

UNIVERSITÀ DI PISA
FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI
CORSO DI LAUREA IN MATEMATICA

TESI DI LAUREA

Dall'integrale stocastico al problema di Dirichlet

CANDIDATO
Andrea Carpignani

RELATORE
Prof. Paolo Acquistapace

CONTRORELATORE
Dott. Marzia De Donno

Anno accademico 2003/2004

Sommario

Introduzione	iii
I Elementi di teoria generale dei processi	1
1. La nozione di processo stocastico	1
2. Intervalli stocastici	4
3. Processi progressivamente misurabili	5
4. Processi prevedibili	7
5. Tempi d'arresto	9
II Martingale	15
1. La nozione di martingala	15
2. Martingale a tempi discreti	17
3. Martingale a tempi continui	20
4. Martingale di quadrato integrabile	27
III Il processo di Wiener	29
1. Variabili aleatorie gaussiane reali	29
2. Definizione di processo di Wiener	31
3. Costruzione del processo di Wiener	35
4. Proprietà delle traiettorie di un processo di Wiener	38
5. Il processo di Wiener multidimensionale	41
IV L'integrale stocastico	43
1. Convergenze negli spazi dei processi	43
2. La nozione di semimartingala e d'integrale stocastico	48
3. Approssimazione di un integrale stocastico	53
4. Un esempio di integrazione stocastica: l'integrale di Stieltjes	55
5. Un esempio di integrazione stocastica: l'integrale di Itô	57
6. Variazione quadratica di una semimartingala	62
7. La formula di Itô	67

8. I processi di Itô	68
9. Cenno sull'integrale stocastico multidimensionale	69
V Equazioni differenziali stocastiche	73
1. Definizione e prime proprietà	73
2. Un risultato di esistenza ed unicità	75
3. Unicità debole delle soluzioni	78
VI Processi di diffusione	81
1. Natura markoviana delle soluzioni	81
2. Generatore di una diffusione	88
3. Formula di Feynman-Kač	94
VII Il problema di Dirichlet-Poisson	97
1. Introduzione al problema di Dirichlet-Poisson	97
2. Teoremi di esistenza ed unicità della soluzione	99
3. Costruzione della soluzione per l'operatore di Laplace	104
A Richiami di Analisi e Probabilità	109
1. Spazi di Riesz e classi monotone	109
2. La nozione di speranza condizionale	112
3. La nozione di nucleo	114
5. Il teorema delle contrazioni	117
5. Funzioni armoniche	118
Bibliografia	119

Introduzione

La nozione di integrale stocastico fu introdotta per la prima volta da N. Wiener nel 1923: si trattava dell'integrale stocastico di un processo "deterministico", ossia indipendente dalla variabile ω , rispetto al processo di Wiener (o moto browniano). La teoria fu poi sviluppata da K. Itô in una serie di articoli pubblicati tra il 1944 e il 1951 (si veda [26], [24], [25], [23]). Egli sviluppò la teoria dell'integrazione stocastica rispetto ad un processo di Wiener e introdusse per la prima volta la nozione di equazione differenziale stocastica, che può essere vista come una generalizzazione della nozione di equazione differenziale ordinaria e ne eredita in modo naturale le proprietà. La prima estensione della teoria al di là del caso "browniano" si trova nel classico libro di J.L. Doob sui processi stocastici (si veda [17]): si tratta stavolta di integrazione rispetto ad una martingala Y , per la quale si suppone l'esistenza d'una funzione F crescente, tale che il processo $(Y_t^2 - F(t))$ sia ancora una martingala. Nel 1962 P.-A. Meyer generalizzò, dapprima solo nel caso discreto, la costruzione di Doob, compiendo un passo decisivo per l'estensione della teoria dell'integrazione stocastica rispetto a martingale di quadrato integrabile e martingale locali (completata poi da H. Kunita, S. Watanabe, P.-A. Meyer). Quest'estensione, come del resto tutti gli ulteriori progressi di questa teoria, ha contribuito a sviluppare la "teoria generale" dei processi stocastici, strumento di rara finezza ed eleganza, creato da P.-A. Meyer, in collaborazione con illustri Autori quali ad esempio C. Dellacherie.

L'apice della teoria dell'integrazione stocastica è stato raggiunto con l'introduzione, da parte di P.-A. Meyer, della nozione di *semimartingala* e del suo studio da parte di numerosi probabilisti (quali C. Dellacherie, H. Kunita, S. Watanabe, per citarne solo alcuni). La sua definizione più classica può essere formulata nel modo seguente: si chiama semimartingala un processo che si possa scrivere come somma di un processo di Stieltjes e di una martingala locale. Più tardi (1984) G. Letta ha formulato una definizione alternativa introducendo la nozione di semimartingala in modo "assiomatico". Che una semimartingala secondo questa definizione sia una semimartingala "classica" si riconosce facilmente (si veda [31], [39]); l'implicazione inversa, e quindi l'equivalenza delle due definizioni, segue da un teorema dovuto a C. Dellacherie (si veda [30]).

Lo scopo di questa tesi è quello di introdurre la nozione di semimartingala (alla maniera "assiomatica") e di studiarne le proprietà elementari, per poi vederne le

più moderne applicazioni al caso delle equazioni differenziali stocastiche, dando, per queste ultime, una definizione rigorosa, illustrandone le proprietà fondamentali e mostrandone i collegamenti con la teoria delle equazioni differenziali ordinarie. Si descrive infine, a mo' di esempio, un'importante applicazione della teoria del calcolo stocastico alle equazioni alle derivate parziali di tipo ellittico, dimostrando che, nel caso in cui un'equazione ellittica abbia soluzione classica, allora questa può essere scritta esplicitamente.

Questa tesi nasce anche con l'ambizione, forse troppo pretenziosa, di colmare il baratro che spesso esiste tra probabilisti e analisti, relativamente sia al linguaggio, sia – più in generale – alla sensibilità, cercando concretamente di rendersi leggibile a qualunque esperto di Analisi disposto a faticare un po'. Il linguaggio e le notazioni che adotteremo saranno quelle classiche adottate da C. Dellacherie e P.-A. Meyer nella celebre collana *Probabilités et potentiel* (si veda [14], [15], [16]), e, soprattutto, quelle adottate dal mio maestro Giorgio Letta in [32], come ad esempio l'utilizzo del termine bourbakista “tribù” in luogo dell'abituale termine σ -algebra, il quale, a stretto rigore di termini, è impreciso e non dovrebbe essere utilizzato (si veda a questo proposito [22]). Entrando nel dettaglio delle notazioni, se X è una variabile aleatoria reale sullo spazio probabilizzato (Ω, \mathcal{A}, P) , converremo di denotare con $P[X]$ la speranza rispetto alla misura P (quando abbia senso), piuttosto che con il classico simbolo $E[X]$, allo scopo di evidenziarne la dipendenza dalla misura di probabilità. Per lo stesso motivo, se \mathcal{F} è una sottotribù di \mathcal{A} , e X è una variabile aleatoria integrabile (secondo P), la speranza condizionale di X dato \mathcal{F} verrà indicata con $P[X | \mathcal{F}]$, anziché con l'usuale notazione $E[X | \mathcal{F}]$. Talvolta, in luogo di $\int f d\mu$, si converrà di denotare con $\langle \mu, f \rangle$ l'integrale di una funzione reale f , misurabile su (E, \mathcal{E}, μ) .

Vediamo esplicitamente il contenuto dei varî capitoli.

Allo scopo di rendere la trattazione autocontenuta, abbiamo riportato nel capitolo I, per comodità del lettore, le linee principali della teoria dei processi stocastici; la terminologia e le notazioni sono fatte alla maniera del *Quaderno* di G. Letta (si veda [31]). Le sole nozioni che si danno per scontate sono quelle tradizionalmente oggetto dei corsi di probabilità elementare (contenute in [32]).

Il capitolo II riguarda le nozioni di martingala e sottomartingala, e lo studio delle proprietà necessarie ad introdurre il processo di Wiener e le semimartingale. A questo scopo, molte delle dimostrazioni sono state riadattate dai testi di C. Dellacherie e P.-A. Meyer [15] e G. Letta [31]. Per non appesantire troppo la trattazione abbiamo deciso di trascurare varie questioni classiche, che però sono ampiamente discusse nei due testi citati.

Il capitolo III si occupa del processo di Wiener e dello studio delle sue principali proprietà. Innanzitutto si pone l'attenzione sulla nozione di “misura gaussiana” e si studiano le variabili aleatorie aventi questa misura come propria legge. Viene poi introdotta la definizione di processo di Wiener e vengono analizzate le principali proprietà di questo processo, tra cui il fatto di essere una martingala. Poiché su questo processo si basa tutta la trattazione delle equazioni differenziali stocastiche,

abbiamo ritenuto opportuno dimostrare l'effettiva esistenza del processo di Wiener (la quale richiede l'introduzione del celebre teorema di hölderianità di Kolmogorov) dedicandovi un'intera sezione. Infine si descrivono alcune proprietà fini delle traiettorie di questo processo.

Nel capitolo IV introduciamo (alla maniera di [31]) lo spazio dei processi regolari (che alcuni Autori chiamano *càdlàg*) e definiamo su di esso una nozione di convergenza, mostrando che si tratta di uno spazio metrico completo. Studiamo poi lo spazio dei processi prevedibili e, anche in questo caso, andiamo a dare, su di esso, una nozione di convergenza. Introduciamo poi, seguendo il testo di G. Letta [31], la nozione di semimartingala e di integrale stocastico nella sua forma più generale, per poi studiarne le principali proprietà, con la differenza, rispetto a [31], che qui vengono trattati processi non necessariamente nulli in 0 (questo perché le soluzioni delle equazioni differenziali stocastiche, che ragionevolmente dovrebbero essere semimartingale, in generale non sono nulle in 0). Introduciamo, successivamente, la nozione di variazione quadratica di una semimartingala, e la celebre formula di Itô che generalizza il “teorema fondamentale del calcolo integrale” nel caso dell'integrale stocastico, e che fu introdotta per la prima volta da K. Itô per il processo di Wiener, e successivamente estesa da H. Kunita e S. Watanabe, C. Doléans e P.-A. Meyer nel caso generale. Il capitolo si conclude con l'introduzione dei cosiddetti processi di Itô.

Il capitolo V tratta le equazioni differenziali stocastiche introducendone la definizione rigorosa e dimostrando un celebre teorema di esistenza ed unicità della soluzione che generalizza in modo naturale il classico teorema di Cauchy-Lipschitz, sull'esistenza ed unicità della soluzione delle equazioni differenziali ordinarie. Si pone, in particolare, l'attenzione alla dipendenza dal dato iniziale che anch'essa generalizza quella nota per le equazioni differenziali ordinarie. La discussione è nello stesso spirito del testo di B. Øksendal (si veda [37]) cercando però di rendere la trattazione più rigorosa. Si considera anche il problema dell'unicità debole della soluzione di un'equazione differenziale stocastica.

Nel capitolo VI vengono studiate le proprietà delle soluzioni di una classe molto ampia di soluzioni di equazioni differenziali stocastiche, le cosiddette “diffusioni”. Si introduce il “semigruppato di transizione” e si osserva che ogni diffusione è completamente determinata da un nucleo markoviano (detto *realizzazione canonica*) che dipende soltanto dai coefficienti b, σ che compaiono nella definizione della diffusione. Questo tipo di soluzioni è un caso speciale di una teoria più generale: la teoria dei *processi di Markov*, nella quale però non ci addentreremo; il lettore desideroso di approfondire questa nozione può consultare il celebre *Lecture Notes* di P.-A. Meyer (si veda [36]). Tramite il semigruppato di transizione è anche possibile associare, ad ogni diffusione, un operatore (detto *generatore infinitesimale*) del quale studieremo alcune proprietà. Il capitolo si conclude con una prima applicazione alle equazioni differenziali di tipo stocastico, scrivendo la soluzione esplicita di una certa classe di equazioni differenziali di tipo parabolico tramite una formula che prende il nome di *formula di Feynman-Kač* (sviluppata ad esempio in [13]). Anche in questo capitolo seguiamo lo spirito del già citato testo di B. Øksendal, ma cercando di esaltare il

senso probabilistico della teoria, così come avviene nei celebri testi di I. Karatzas e S. Shreve (si veda [28]) e di P. Baldi (si veda [4]).

Il capitolo VII, come promesso, analizza il problema di Dirichlet-Poisson

$$\begin{cases} Lu = -g & \text{su } D, \\ u(x) = \phi(x) & \text{per ogni } x \in \partial D, \end{cases}$$

attraverso lo studio della diffusione avente l'operatore L come generatore infinitesimale. Si dimostra infatti un teorema secondo il quale, se un problema ellittico ha soluzione di classe \mathcal{C}^2 , allora essa è unica e si scrive esplicitamente tramite una formula di rappresentazione che dipende soltanto dalla realizzazione canonica della diffusione. Il problema dell'esistenza della soluzione è demandato, nel caso generale, ad un classico risultato di esistenza (si veda D. Gilbarg, N.S. Trudinger [20]). L'interpretazione di tale soluzione, nel caso del problema di Dirichlet, è la seguente: *la soluzione $u(x)$ coincide con il valore atteso di g nel primo istante d'uscita dal dominio D della diffusione che parte da x .*

Analizzando il problema di Poisson, si scopre poi che la soluzione può essere scritta come la trasformata del termine noto mediante un nucleo (detto *nucleo di Green*) che si può scrivere esplicitamente in termini della realizzazione canonica. Nel caso poi dell'operatore di Laplace, il nucleo di Green si scrive come *nucleo densità*, ossia nella forma

$$G(x, A) = \int_A G(x, y) dy$$

dove la funzione G , di \mathbb{R}^2 in \mathbb{R} , coincide con la classica funzione di Green associata al problema. Questo permette anche di scrivere in modo esplicito (in termini della realizzazione canonica del processo di Wiener) la funzione di Green.

Nel caso specifico del problema di Dirichlet per l'equazione di Laplace, è possibile provare l'esistenza della soluzione senza ricorrere ai teoremi classici di esistenza, non appena il dominio verifica la cosiddetta "proprietà del cono" (per ulteriori approfondimenti su questo aspetto si può consultare K.L. Chung [11]).

Tante sono le persone che dovrei ringraziare per i loro consigli, le lunghe chiacchierate e l'infinita pazienza. Mi limiterò a citarne solo alcune sperando nella clemenza delle altre che comunque sono ben consapevoli della stima che provo per loro. La prima persona è sicuramente Paolo Acquistapace: non dimenticherò mai tutto l'aiuto e la pazienza che ha avuto con me, prima come studente di Analisi e poi come laureando; in più di un'occasione la sua lucidità è stata indispensabile. È stato un vero onore per me essere stato suo studente. Desidero poi ringraziare Marzia De Donno: senza il suo aiuto non sarei mai riuscito a completare questa tesi.

Vorrei ringraziare, infine, la persona che, con la sua precisione ed il suo ineccepibile gusto estetico, mi ha fatto innamorare della matematica più di quanto già lo fossi: queste poche parole di ringraziamento non rendono certo giustizia all'importante ruolo che ha avuto, nello svolgimento di questa tesi e nella mia formazione di matematico, il mio indiscusso maestro Giorgio Letta al quale, in un certo senso, questa tesi è dedicata.

1.1 La nozione di processo stocastico

Dati uno spazio probabilizzato (Ω, \mathcal{A}, P) , uno spazio misurabile (E, \mathcal{E}) e un insieme ordinato I , un *processo* a valori in (E, \mathcal{E}) , avente I come insieme dei tempi si definisce tradizionalmente come una famiglia $(X_t)_{t \in I}$ di variabili aleatorie su (Ω, \mathcal{A}, P) a valori in (E, \mathcal{E}) e avente I come insieme degli indici. Nella teoria generale si preferisce pensare un processo piuttosto come un'applicazione X da $I \times \Omega$ a E , tale che, per ciascun t , l'applicazione parziale $\omega \mapsto X(t, \omega)$ (denotata con X_t) risulti una variabile aleatoria. Le applicazioni parziali $t \mapsto X(t, \omega)$ vengono dette le *traiettorie* del processo.

(1.1.1) Definizione. Per due processi X e Y definiti sul medesimo spazio, si dice che Y è una *modificazione* di X se, per ogni fissato istante t , le due variabili aleatorie X_t e Y_t coincidono quasi certamente, cioè fuori da un insieme di probabilità nulla (dipendente da t).

Nella teoria generale è utile considerare, accanto a questa nozione, anche un concetto di equivalenza molto più restrittivo:

(1.1.2) Definizione. Due processi X e Y definiti sul medesimo spazio si dicono *indistinguibili* se, per quasi ogni ω , le due traiettorie $X(\cdot, \omega)$ e $Y(\cdot, \omega)$ sono identiche.

Ciò significa in altri termini che, nello spazio $I \times \Omega$, l'insieme $\{X \neq Y\}$ è *evanescente*, nel senso che la sua proiezione su Ω è contenuta in un insieme di probabilità nulla.

(1.1.3) Esempio. Per provare che le due relazioni d'equivalenza introdotte non coincidono, si consideri lo spazio probabilizzato costituito dall'intervallo $\Omega = [0, 1]$ munito della propria tribù boreliana \mathcal{A} e della misura di Borel-Lebesgue P . Su tale spazio si consideri il processo X , avente $[0, 1]$ come insieme dei tempi, definito da

$$X(t, \omega) = \begin{cases} t & \text{se } \omega = t \\ 0 & \text{altrimenti.} \end{cases}$$

È chiaro che X sia una modificazione del processo nullo Y (definito da $Y(t, \omega) = 0$), essendo ogni variabile aleatoria X_t equivalente modulo P alla costante 0. D'altra parte, per nessun ω diverso da 0 in Ω , la traiettoria $X(\cdot, \omega)$ coincide con la costante 0 e quindi X non è indistinguibile dal processo nullo Y .

(1.1.4) Definizione. Una *filtrazione* sullo spazio probabilizzato (Ω, \mathcal{A}, P) , avente I come insieme dei tempi, è una famiglia $\mathcal{F} = (\mathcal{F}_t)_{t \in I}$ di sottotribù di \mathcal{A} , avente I come insieme degli indici e tale che si abbia $\mathcal{F}_s \subseteq \mathcal{F}_t$ per ciascuna coppia s, t di elementi di I , con $s \leq t$. Gli elementi di \mathcal{F}_t si dicono *anteriori all'istante t* (relativamente alla filtrazione \mathcal{F}).

(1.1.5) Definizione. Supponiamo fissata una filtrazione $(\mathcal{F}_t)_{t \in I}$ sullo spazio probabilizzato (Ω, \mathcal{A}, P) , un processo X avente come insieme dei tempi una parte J di I , si dice *adattato a \mathcal{F}* se, per ogni elemento t di J , la variabile aleatoria X_t risulta \mathcal{F}_t -misurabile.

Dato un processo X sullo spazio (Ω, \mathcal{A}, P) , per ciascun t , si denoti con \mathcal{G}_t la tribù generata dal blocco delle variabili aleatorie X_s con $s \leq t$. La famiglia $\mathcal{G} = (\mathcal{G}_t)_{t \in I}$ risulta, dunque, una filtrazione su (Ω, \mathcal{A}, P) detta *filtrazione naturale di X* . È chiaro che, affinché X sia un processo adattato ad una filtrazione \mathcal{F} , è necessario e sufficiente che, per ciascun indice t , si abbia $\mathcal{G}_t \subseteq \mathcal{F}_t$.

Nel seguito supponiamo fissata una filtrazione $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ avente come insieme dei tempi l'intervallo $\mathbb{R}_+ = [0, \infty)$. I processi considerati nel seguito saranno per semplicità processi sullo spazio probabilizzato (Ω, \mathcal{A}, P) a valori nella retta reale \mathbb{R} munita della propria tribù boreliana $\mathcal{B}(\mathbb{R})$, ossia processi *reali*, aventi come insieme dei tempi l'intervallo \mathbb{R}_+ ; si potrebbe tuttavia considerare processi a valori in un qualsiasi spazio di Hilbert separabile, munito della propria tribù boreliana, senza modificare nessun risultato. Si designa con \mathcal{F}^+ la filtrazione definita, a partire da \mathcal{F} , nel modo seguente:

$$(1.1.6) \quad \mathcal{F}_t^+ = \bigcap_{u > t} \mathcal{F}_u, \quad \mathcal{F}_\infty^+ = \mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t.$$

(1.1.7) Definizione. Si dice che la filtrazione $(\mathcal{F}_t)_{t \geq 0}$ verifica le *condizioni abituali* se possiede le due proprietà seguenti:

(a) \mathcal{F} coincide con la filtrazione \mathcal{F}^+ definita in (1.1.6) (cosa che si esprime dicendo che \mathcal{F} è *continua a destra*).

(b) Ogni elemento di \mathcal{A} , trascurabile secondo P , appartiene alla tribù \mathcal{F}_0 (cosa che si esprime dicendo che \mathcal{F}_0 è *satura*).

(1.1.8) Definizione. Una funzione reale f , definita in una parte A di \mathbb{R} è detta *regolare a destra* se, per ogni elemento s di \mathbb{R} tale che l'insieme $A \cap (s, \infty)$ ammetta s come punto di accumulazione, il limite

$$\lim_{t \downarrow s} f(t)$$

esiste finito. Si definisce in modo analogo la nozione di funzione *regolare a sinistra*. Quando una funzione f sia contemporaneamente regolare a destra e a sinistra si

dice semplicemente che essa è *regolare*. Un processo reale X la cui traiettorie siano tutte regolari si dice un *processo regolare*. In modo analogo si possono definire le nozioni di *processo continuo*, *continuo a destra* etc.

Un processo X che abbia $X_0 = 0$ sarà detto *nullo all'istante 0* (o *nullo in 0*). Un processo X indistinguibile dal processo nullo (cioè tale che l'insieme $\{X \neq 0\}$ sia evanescente) sarà detto *evanescente*.

(1.1.9) Proposizione. *Se X e Y sono processi continui a destra, affinché Y sia indistinguibile da X , è (necessario e) sufficiente che Y sia una modificazione di X .*

Dimostrazione. Si supponga che Y sia una modificazione di X e si ponga

$$A = \bigcup_{t \in \mathbb{Q}_+} \{X_t \neq Y_t\}.$$

L'insieme A è allora trascurabile. Inoltre, per ogni elemento ω di A^c , le due traiettorie $X(\cdot, \omega)$ e $Y(\cdot, \omega)$ (continue a destra e aventi la stessa restrizione a \mathbb{Q}_+) sono identiche.

(1.1.10) Notazioni. Sia X un processo reale avente \mathbb{R}_+ come insieme dei tempi.

(a) Se X è regolare a destra, si denota con X_+ il processo (avente \mathbb{R}_+ come insieme dei tempi) definito come

$$X_+(t, \omega) = \lim_{s \downarrow t} X(s, \omega).$$

(b) Se X è regolare a sinistra, si denota con X_- il processo definito come

$$X_-(0, \omega) = 0 \quad \text{e} \quad X_-(t, \omega) = \lim_{s \uparrow t} X(s, \omega).$$

Con le notazioni appena definite è immediato provare la proposizione seguente.

(1.1.11) Proposizione. *Sia X un processo avente \mathbb{R}_+ come insieme dei tempi.*

(a) *Se X è regolare a destra e adattato, allora il processo X_+ è continuo a destra e adattato a \mathcal{F}^+ .*

(b) *Se X è regolare a sinistra e adattato a \mathcal{F}^+ , allora il processo X_- è continuo a sinistra e adattato a \mathcal{F} .*

1.2 Intervalli stocastici

La seguente notazione risulterà indispensabile: siano S e T due applicazioni da Ω a $[0, \infty]$, con $S \leq T$, si ponga

$$\begin{aligned} \llbracket S, T \rrbracket &= \{(t, \omega) \in \mathbb{R}_+ \times \Omega : S(\omega) \leq t \leq T(\omega)\}, \\ \llbracket S, T \llbracket &= \{(t, \omega) \in \mathbb{R}_+ \times \Omega : S(\omega) \leq t < T(\omega)\}, \\ \rrbracket S, T \rrbracket &= \{(t, \omega) \in \mathbb{R}_+ \times \Omega : S(\omega) < t \leq T(\omega)\}, \\ \rrbracket S, T \llbracket &= \{(t, \omega) \in \mathbb{R}_+ \times \Omega : S(\omega) < t < T(\omega)\}. \end{aligned}$$

Ciascuna di queste parti di $\mathbb{R}_+ \times \Omega$ è detta un *intervallo stocastico*. In particolare, l'intervallo stocastico $\llbracket T, T \rrbracket$ si denota più semplicemente con $\llbracket T \rrbracket$ e si chiama il *grafico* di T : per definizione, infatti, si ha

$$\llbracket T \rrbracket = \{(t, \omega) \in \mathbb{R}_+ \times \Omega : T(\omega) = t\}.$$

È chiaro che, data una coppia s, t di numeri reali positivi, con $s \leq t$, gli intervalli stocastici $\llbracket s, t \rrbracket$, $\llbracket s, t \llbracket$, $\rrbracket s, t \rrbracket$, $\rrbracket s, t \llbracket$ si riducono rispettivamente ai seguenti sottoinsiemi di $\mathbb{R}_+ \times \Omega$:

$$[s, t] \times \Omega, \quad [s, t[\times \Omega, \quad]s, t] \times \Omega, \quad]s, t[\times \Omega.$$

Sia data un'applicazione T da Ω a $[0, \infty]$ e sia A una parte di Ω . Si denoti con T_A l'applicazione da Ω a $[0, \infty]$ definita come

$$(1.2.1) \quad T_A(\omega) = \begin{cases} T(\omega) & \text{se } \omega \in A \\ \infty & \text{altrimenti.} \end{cases}$$

È chiaro che il grafico di T_A è contenuto nel grafico di T . Precisamente, si ha:

$$\llbracket T_A \rrbracket = \llbracket T \rrbracket \cap (\mathbb{R}_+ \times A).$$

La proposizione seguente (di dimostrazione immediata) sarà utile nel seguito.

(1.2.2) Proposizione. *Sia data una famiglia non vuota $(T_i)_{i \in J}$ d'applicazioni da Ω a $[0, \infty]$. Si ponga:*

$$S = \inf_i T_i, \quad U = \sup_i T_i.$$

Si ha, allora:

$$(1.2.3) \quad \llbracket S, \infty \rrbracket = \bigcup_i \llbracket T_i, \infty \rrbracket, \quad \llbracket U, \infty \llbracket = \bigcap_i \llbracket T_i, \infty \llbracket.$$

In particolare, se J è un insieme finito,

$$(1.2.4) \quad \llbracket S, \infty \rrbracket = \bigcup_i \llbracket T_i, \infty \rrbracket, \quad \llbracket U, \infty \llbracket = \bigcap_i \llbracket T_i, \infty \llbracket.$$

1.3 Processi progressivamente misurabili

Il considerare un processo come funzione della coppia (t, ω) suggerisce in modo spontaneo l'idea di studiarne la misurabilità rispetto a opportune tribù fissate sullo spazio prodotto $\mathbb{R}_+ \times \Omega$. Tra queste tribù la prima che interviene è, naturalmente, la tribù $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{A}$: un processo misurabile rispetto a questa tribù viene detto semplicemente *misurabile*.

La presenza di una filtrazione permette di considerare, per un processo, accanto alla semplice misurabilità, anche una nozione più restrittiva:

(1.3.1) Definizione. Sia \mathcal{F} una filtrazione avente come insieme dei tempi \mathbb{R}_+ , un processo reale X , avente anch'esso \mathbb{R}_+ come insieme dei tempi, si dice *progressivamente misurabile* (o *progressivo*) se, per ogni elemento t di \mathbb{R}_+ , la restrizione di X a $\llbracket 0, t \rrbracket$ è misurabile rispetto alla tribù $\mathcal{B}(\llbracket 0, t \rrbracket) \otimes \mathcal{F}_t$. In modo analogo, si dice che il processo X è progressivamente misurabile *in senso largo* se, per ogni elemento t in \mathbb{R}_+ diverso da 0, la restrizione di X a $\llbracket 0, t \rrbracket$ è misurabile rispetto alla tribù $\mathcal{B}(\llbracket 0, t \rrbracket) \otimes \mathcal{F}_t$.

È chiaro che le due nozioni di misurabilità progressiva appena definite dipendono dalla filtrazione scelta: esse si riconducono alla nozione di misurabilità nel caso in cui \mathcal{F} coincida con la filtrazione “deterministica”, cioè la filtrazione costante definita dalla relazione $\mathcal{F}_t = \mathcal{A}$ per ciascun t in \mathbb{R}_+ .

La proposizione seguente mette in relazione le due nozioni di misurabilità progressiva appena introdotte.

(1.3.2) Proposizione. Assegnata una filtrazione \mathcal{F} , sia X un processo reale su (Ω, \mathcal{A}, P) avente \mathbb{R}_+ come insieme dei tempi. Le condizioni che seguono sono, allora, equivalenti:

- (a) X è progressivo.
- (b) X è progressivo in senso largo e adattato.

Dimostrazione. Si fissi un numero reale positivo t e si consideri la restrizione di X all'insieme $\llbracket 0, t \rrbracket = \llbracket 0, t \rrbracket \cup \llbracket t \rrbracket$. Affinché tale restrizione sia misurabile rispetto alla tribù $\mathcal{B}(\llbracket 0, t \rrbracket) \otimes \mathcal{F}_t$, occorre e basta che siano verificate le due condizioni seguenti:

- (1) se t è strettamente maggiore di zero, la restrizione di X a $\llbracket 0, t \rrbracket$ è misurabile rispetto alla traccia di $\mathcal{B}(\llbracket 0, t \rrbracket) \otimes \mathcal{F}_t$ su $\llbracket 0, t \rrbracket$, cioè rispetto alla tribù $\mathcal{B}(\llbracket 0, t \rrbracket) \otimes \mathcal{F}_t$;
- (2) la restrizione di X a $\llbracket t \rrbracket$ è misurabile rispetto alla traccia di $\mathcal{B}(\llbracket 0, t \rrbracket) \otimes \mathcal{F}_t$ su $\llbracket t \rrbracket$, o ciò ch'è lo stesso, la variabile aleatoria X_t è misurabile rispetto a \mathcal{F}_t .

Per concludere basta osservare che le condizioni (1) e (2) equivalgono appunto al fatto che X è progressivo in senso largo e adattato.

Il criterio di misurabilità progressiva appena dimostrato si semplifica ulteriormente sotto l'ipotesi che il processo sia continuo a destra:

(1.3.3) Proposizione. *Assegnata una filtrazione \mathcal{F} , sia X un processo reale su (Ω, \mathcal{A}, P) avente \mathbb{R}_+ come insieme dei tempi e continuo a destra. Le condizioni che seguono sono, allora, equivalenti:*

- (a) X è progressivo.
- (b) X è adattato.

Dimostrazione. Grazie alla proposizione precedente, tutto è ridotto a mostrare che, se X è un processo adattato, allora è progressivamente misurabile in senso largo. Fissato, a questo proposito, un numero reale strettamente positivo t , una volta posto, per ogni coppia n, k d'interi strettamente positivi,

$$X^{n,k}(s, \omega) = \begin{cases} X(k2^{-n}, \omega) & \text{per } (k-1)2^{-n} \leq s < k2^{-n} \\ 0 & \text{altrimenti,} \end{cases}$$

si consideri la successione (X^n) definita, per ogni intero n strettamente positivo, come

$$X^n(t, \omega) = \sum_{k \geq 1, k2^{-n} < t} X^{n,k}(t, \omega).$$

Poiché le traiettorie di X sono continue a destra, la successione (X^n) converge puntualmente verso X su $\llbracket 0, t \rrbracket$. Inoltre, ciascuno dei processi $X^{n,k}$ ha restrizione a $\llbracket 0, t \rrbracket$ misurabile rispetto alla tribù $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$. Lo stesso si può allora dire per il processo X .

Applicando il risultato precedente alla filtrazione deterministica si ottiene il corollario seguente:

(1.3.4) Corollario. *Ogni processo continuo a destra è misurabile.*

La proposizione seguente permette di riportare la nozione di processo progressivo in senso largo a quella di processo progressivo: essa in particolare, mostra che, se una filtrazione è continua a destra, allora le due nozioni coincidono.

(1.3.5) Proposizione. *Assegnata una filtrazione \mathcal{F} , sia X un processo reale su (Ω, \mathcal{A}, P) avente \mathbb{R}_+ come insieme dei tempi e continuo a destra. Le condizioni che seguono sono, allora, equivalenti:*

- (a) X è un processo progressivo in senso largo rispetto alla filtrazione \mathcal{F} .
- (b) X è un processo progressivo rispetto alla filtrazione \mathcal{F}^+ .

Dimostrazione. Per quanto riguarda l'implicazione (a) \Rightarrow (b): si supponga che il processo X sia progressivo in senso largo (rispetto a \mathcal{F} e, di conseguenza, rispetto a \mathcal{F}^+). Applicando la proposizione (1.3.2) alla filtrazione \mathcal{F}^+ basta provare che X è adattato alla filtrazione \mathcal{F}^+ . Fissato, a questo proposito, un numero reale positivo t , per ogni numero reale $u > t$, per ipotesi, la restrizione di X a $\llbracket 0, u \rrbracket$ è misurabile rispetto alla tribù $\mathcal{B}([0, u]) \otimes \mathcal{F}_u$. Ne segue, in particolare, che X_t è misurabile rispetto a \mathcal{F}_u . Poiché questa conclusione vale per ciascun numero reale $u > t$, essa mostra che X_t è misurabile rispetto a \mathcal{F}_t^+ .

Viceversa, per provare l'implicazione (b) \Rightarrow (a): fissato un numero reale strettamente positivo t e scelta una successione crescente (t_n) convergente verso t , si ha:

$$\llbracket 0, t \llbracket = \bigcup_{n \in \mathbb{N}} \llbracket 0, t_n \llbracket.$$

Per l'ipotesi (b), la restrizione di X all'intervallo stocastico $\llbracket 0, t_n \llbracket$ è misurabile rispetto alla tribù $\mathcal{B}([0, t_n]) \otimes \mathcal{F}_{t_n}^+$ e quindi anche rispetto alla tribù $\mathcal{B}([0, t_n]) \otimes \mathcal{F}_t$. Ne segue necessariamente che la restrizione di X all'insieme $\llbracket 0, t \llbracket$ è misurabile rispetto alla tribù $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ e ciò conclude la dimostrazione.

1.4 Processi prevedibili

Si chiama *processo misurabile elementare* il processo ottenuto come la somma di un numero finito di processi della forma

$$(1.4.1) \quad I_{]s, t]} \otimes V,$$

dove s, t sono numeri reali positivi, con $s < t$ e V è una variabile aleatoria reale e limitata.

Data una filtrazione \mathcal{F} su (Ω, \mathcal{A}, P) , avente \mathbb{R}_+ come insieme dei tempi, si chiama *processo prevedibile elementare* il processo ottenuto come la somma di un numero finito di processi della forma (1.4.1), dove s, t sono numeri reali positivi, con $s < t$, e V è una variabile aleatoria reale e limitata, misurabile rispetto a \mathcal{F}_s . Questa nozione si riduce a quella di processo elementare quando \mathcal{F} sia la filtrazione deterministica.

È immediato verificare che ogni processo prevedibile elementare è, in particolare, un processo progressivamente misurabile e, quindi, adattato.

(1.4.2) Definizione. Si chiama *tribù prevedibile* la tribù generata su $\llbracket 0, \infty \llbracket$ da tutti i processi prevedibili elementari, e si chiamano *insiemi prevedibili* tutti gli elementi di tale tribù.

È immediato mostrare che la tribù prevedibile è generata da tutte le parti di $\llbracket 0, \infty \llbracket$ della forma $]s, t] \times A$, con s, t numeri reali positivi, con $s < t$ e A appartenente a \mathcal{F}_s . Tali sottoinsiemi di $\llbracket 0, \infty \llbracket$ vengono detti *rettangoli prevedibili*: essi sono degli intervalli stocastici, potendo scrivere ogni $]s, t] \times A$, attraverso la notazione (1.2.1), nella forma $\llbracket s_A, t_A \llbracket$. Inoltre, è immediato verificare che ogni rettangolo prevedibile è un insieme progressivo, ossia tale è la sua funzione indicatrice.

In particolare, una volta osservato che l'insieme di tutte le parti progressive di $\llbracket 0, \infty \llbracket$ costituisce una tribù, risulta evidente che ogni insieme prevedibile è progressivo, ossia, la tribù prevedibile risulta una sottotribù della tribù progressiva. Ne segue che, *tutte le applicazioni X da $\llbracket 0, \infty \llbracket$ a \mathbb{R} , misurabili rispetto alla tribù prevedibile sono processi progressivi (e quindi adattati) aventi $]0, \infty[$ come insieme dei tempi.*

(1.4.3) Definizione. Un processo X si dice *prevedibile* se X_0 è misurabile rispetto alla tribù \mathcal{F}_0 e se la restrizione di X a $]0, \infty[$ è misurabile rispetto alla tribù prevedibile.

Una classe molto importante di processi prevedibili è fornita dalla proposizione seguente.

(1.4.4) Proposizione. Ogni processo reale X su (Ω, \mathcal{A}, P) , adattato e continuo a sinistra è prevedibile.

Dimostrazione. Poiché X è adattato, basta dimostrare l'asserzione per processi nulli in 0. Si ponga, per ogni intero positivo n , $X^n = \sum_{k \geq 1} X^{n,k}$ dove $X^{n,k}$ designa il processo definito come

$$X^{n,k}(t, \omega) = \begin{cases} X((k-1)2^{-n}, \omega) & \text{se } (k-1)2^{-n} < t \leq k2^{-n} \\ 0 & \text{altrimenti.} \end{cases}$$

Poiché ogni traiettoria di X è continua a sinistra, la successione (X^n) converge puntualmente verso X . Per concludere basta osservare che ogni processo elementare X^n è prevedibile, quindi tale dev'essere il processo X .

Osserviamo che è possibile caratterizzare la tribù prevedibile come quella generata su $]0, \infty[$ dalla restrizione a $]0, \infty[$ dei processi del tipo descritto nella proposizione precedente perché la classe di questi processi contiene le funzioni indicatrici dei rettangoli prevedibili. Inoltre, affinché un processo continuo a sinistra, avente $]0, \infty[$ come insieme dei tempi, sia adattato a \mathcal{F} occorre e basta che esso sia adattato a \mathcal{F}^+ (si veda a tal proposito l'asserzione (b) della proposizione (1.1.11)) e quindi la tribù prevedibile relativa alla filtrazione \mathcal{F}^+ coincide con quella relativa alla filtrazione \mathcal{F} .

Una volta fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) , accanto alla filtrazione \mathcal{F} può essere utile considerare una filtrazione $\mathcal{G} = (\mathcal{G}_t)_{t \in [0, \infty]}$ più fine (detta *ingrossamento* di \mathcal{F}). Si può considerare, per esempio, la filtrazione \mathcal{G} definita, per ciascun indice t , come $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{N}$, dove \mathcal{N} denota la tribù costituita da tutti gli eventi degeneri secondo P . Tale filtrazione verifica evidentemente la seconda delle "condizioni abituali" (si veda (1.1.7)). Dunque, la filtrazione \mathcal{G}^+ verifica entrambe le condizioni abituali: tale filtrazione prende il nome di *ingrossamento abituale* di \mathcal{F} .

La seguente proposizione lega la nozione di prevedibilità relativa ad un'assegnata filtrazione \mathcal{F} a quella relativa al proprio ingrossamento abituale.

(1.4.5) Proposizione. Su uno spazio probabilizzato (Ω, \mathcal{A}, P) , sia \mathcal{F} una fissata filtrazione avente \mathbb{R}_+ come insieme dei tempi. Si denoti con \mathcal{G} la filtrazione definita, per ciascun t , come $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{N}$, dove \mathcal{N} denota la tribù degli eventi degeneri secondo P . Allora:

(a) Per ogni rettangolo prevedibile G relativo alla filtrazione \mathcal{G} esiste un rettangolo prevedibile G' relativo alla filtrazione \mathcal{F} tale che la differenza simmetrica $G \Delta G'$ sia evanescente.

(b) Per ogni processo limitato H , prevedibile rispetto a \mathcal{G} (o ciò ch'è lo stesso rispetto all'ingrossamento abituale \mathcal{G}^+ di \mathcal{F}), esiste un processo limitato H' , prevedibile rispetto a \mathcal{F} e indistinguibile da H .

Dimostrazione. La prima condizione è conseguenza immediata della definizione di \mathcal{G} . L'altra condizione si ottiene, attraverso il teorema delle classi monotone, osservando che la classe costituita da tutti i processi H limitati e prevedibili rispetto a \mathcal{G} tali che la desiderata proprietà sia soddisfatta è uno spazio vettoriale monotono contenente la classe delle indicatori dei rettangoli prevedibili.

1.5 Tempi d'arresto

Sia fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, una filtrazione \mathcal{F} avente \mathbb{R}_+ come insieme dei tempi.

(1.5.1) Definizione. Una funzione T da Ω a valori in $[0, \infty]$ si dice un *tempo d'arresto* se l'intervallo stocastico $\llbracket T, \infty \rrbracket$ è adattato (cioè se tale è la sua indicatrice), ovvero se, per ogni numero reale positivo t , si ha:

$$\{T \leq t\} \in \mathcal{F}_t.$$

Se T è un tempo d'arresto, la funzione indicatrice dell'intervallo stocastico $\llbracket T, \infty \rrbracket$ è un processo adattato e continuo a destra; quindi, per la proposizione (1.3.3), è un processo progressivo. Dunque la definizione di tempo d'arresto non cambia se si richiede che T sia una variabile aleatoria a valori in $[0, \infty]$ tale che l'intervallo stocastico $\llbracket T, \infty \rrbracket$ sia progressivo. Quest'osservazione suggerisce la definizione seguente:

(1.5.2) Definizione. Si chiama *tempo d'arresto largo* una funzione T da Ω a valori in $[0, \infty]$ tale che l'intervallo stocastico $\llbracket T, \infty \rrbracket$ sia progressivo in senso largo, ossia (per la proposizione (1.3.5)) se T è un tempo d'arresto per la filtrazione \mathcal{F}^+ .

È chiaro che, quando la filtrazione \mathcal{F} è continua a destra, allora la nozione di tempo d'arresto largo coincide con quella di tempo d'arresto. In generale, però, le due nozioni appena introdotte si possono caratterizzare anche nel modo seguente:

(1.5.3) Proposizione. Sia T una funzione da Ω a valori in $[0, \infty]$. Le condizioni che seguono sono, allora, equivalenti:

- (a) T è un tempo d'arresto largo (ossia un tempo d'arresto per la filtrazione \mathcal{F}^+).
- (b) L'intervallo stocastico $\llbracket T, \infty \rrbracket$ è prevedibile.
- (c) L'intervallo stocastico $\llbracket T, \infty \rrbracket$ è adattato (cioè, per ciascun numero reale strettamente positivo t , si ha $\{T < t\} \in \mathcal{F}_t$).

Dimostrazione. Cominciamo con l'implicazione (a) \Rightarrow (b): si ponga $X = I_{\llbracket T, \infty \rrbracket}$. L'ipotesi (a) significa che X è un processo adattato a \mathcal{F}^+ : attraverso la proposizione

(1.1.11), ne segue che il processo $X_- = I_{\llbracket T, \infty \rrbracket}$ è continuo a sinistra e adattato e, dunque, prevedibile per la proposizione (1.4.4). L'implicazione (b) \Rightarrow (c) segue dal fatto che tutti gli insiemi prevedibili sono adattati. Infine, per quanto riguarda l'implicazione (c) \Rightarrow (a): si ponga $Y = I_{\llbracket T, \infty \rrbracket}$. L'ipotesi (c) significa che Y è un processo adattato e, dalla proposizione (1.1.11), ne segue che il processo $Y_+ = I_{\llbracket T, \infty \rrbracket}$ è adattato a \mathcal{F}^+ .

(1.5.4) Proposizione. *Affinché una funzione T , da Ω a valori in $[0, \infty]$, sia un tempo d'arresto largo, è necessario e sufficiente che per ogni numero reale positivo t , la funzione $T \wedge t$ sia \mathcal{F}_t -misurabile.*

Dimostrazione. La condizione è chiaramente necessaria perché, se T è un tempo d'arresto largo, per ogni coppia s, t di numeri reali positivi, con $s \leq t$, si ha:

$$\{T \wedge t \leq s\} = \{T < s\} \in \mathcal{F}_s \subset \mathcal{F}_t.$$

D'altra parte la condizione è anche sufficiente: infatti, per ogni numero reale positivo t , si ha

$$\{T < t\} = \{T \wedge t < t\} \in \mathcal{F}_t.$$

Le proprietà principali di stabilità della classe dei tempi d'arresto sono riassunte nella proposizione seguente:

(1.5.5) Proposizione. (a) *La classe dei tempi d'arresto è stabile per le operazioni di inviluppo superiore numerabile e inviluppo inferiore finito.*

(b) *La classe dei tempi d'arresto larghi è stabile per le operazioni di inviluppo superiore numerabile e inviluppo inferiore numerabile.*

Dimostrazione. (a) Nelle ipotesi della proposizione (1.2.2), si supponga che T_i siano tempi d'arresto e che I sia numerabile. La seconda relazione di (1.2.3) mostra allora che la funzione $U = \sup_i T_i$ è un tempo d'arresto. Nel caso in cui I sia finito, la prima relazione di (1.2.4) mostra che la funzione $S = \inf_i T_i$ è un tempo d'arresto.

(b) Nelle ipotesi della proposizione (1.2.2), si supponga, adesso, che T_i siano tempi d'arresto larghi e che I sia numerabile. La prima relazione di (1.2.3) mostra allora che la funzione $S = \inf_i T_i$ è un tempo d'arresto largo. Inoltre, applicando alla filtrazione \mathcal{F}^+ la parte (a) appena provata, si vede che $U = \sup_i T_i$ è un tempo d'arresto relativo alla filtrazione \mathcal{F}^+ .

(1.5.6) Proposizione. *Sia T un tempo d'arresto largo. Allora esiste una successione decrescente di tempi d'arresto discreti tali che si abbia $T = \inf_n T_n$ e, per ogni intero positivo n , $\{T < T_n\} = \{T < \infty\}$.*

Dimostrazione. Per ogni intero strettamente positivo n , si denoti con D_n l'insieme costituito da tutti i numeri diadici di stadio n -esimo (cioè della forma $k2^{-n}$ al variare di k in \mathbb{N}) e si ponga (con la convenzione che $\inf \emptyset = \infty$):

$$T_n(\omega) = \inf\{s \in D_n : T(\omega) < s\}.$$

La funzione T_n così definita è certamente una variabile aleatoria discreta (l'insieme dei suoi valori è contenuto in $D_n \cup \{\infty\}$). Inoltre, per ogni n , T_n è minorata da T e inoltre $T < T_n \leq T + 2^{-n}$ su $\{T < \infty\}$. Infine, per ogni numero reale positivo t , si ha:

$$\{T_n \leq t\} = \bigcup_{s \in D_n, s \leq t} \{T < s\} \in \mathcal{F}_t,$$

che prova che T_n è un tempo d'arresto.

(1.5.7) Corollario. *Nelle stesse ipotesi della proposizione (1.5.6), si supponga per giunta che \mathcal{F} è una filtrazione continua a destra. Per ogni tempo d'arresto T esiste allora una successione decrescente di tempi d'arresto discreti (T_n) , tale che si abbia $\inf_n T_n = T$.*

(1.5.8) Proposizione. *Sia T un tempo d'arresto. Il grafico di T è allora un insieme progressivo. Inoltre, ogni parte progressiva di $\llbracket T \rrbracket$ è a sua volta il grafico di un tempo d'arresto.*

Dimostrazione. È chiaro che si ha $\llbracket T \rrbracket = \llbracket T, \infty \rrbracket \cap (\llbracket T, \infty \rrbracket)^c$. D'altra parte l'insieme $\llbracket T, \infty \rrbracket$ è progressivo e l'insieme $\llbracket T, \infty \rrbracket$ è prevedibile e, dunque, progressivo. Ne segue necessariamente che $\llbracket T \rrbracket$ dev'essere esso stesso progressivo.

Sia, adesso, G una parte progressiva di $\llbracket T \rrbracket$. Si denoti con A la sua proiezione su Ω . Si ha allora, con la notazione (1.2.1), $G = \llbracket T_A \rrbracket$. Basta, dunque, dimostrare che T_A è un tempo d'arresto. Difatti, per ogni numero reale positivo t ,

$$\begin{aligned} (1.5.9) \quad \{T_A \leq t\} &= A \cap \{T \leq t\} = \{T \leq t\} \cap \{\omega : (T(\omega), \omega) \in G\} \\ &= \{T \leq t\} \cap \{\omega : (T(\omega) \wedge t, \omega) \in G \cap \llbracket 0, t \rrbracket\}. \end{aligned}$$

Inoltre, l'applicazione $\omega \mapsto (T(\omega) \wedge t, \omega)$, dallo spazio misurabile (Ω, \mathcal{F}_t) allo spazio misurabile $(\llbracket 0, t \rrbracket, \mathcal{B}([0, t]) \otimes \mathcal{F}_t)$, risulta misurabile grazie alla proposizione (1.5.6), e l'insieme $G \cap \llbracket 0, t \rrbracket$ appartiene alla tribù $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ perché G è progressivo. Ne risulta che l'ultimo membro di (1.5.9) è un elemento di \mathcal{F}_t e ciò basta per concludere.

Si supponga fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, si supponga fissata una filtrazione \mathcal{F} avente \mathbb{R}_+ come insieme dei tempi. Sia T un tempo d'arresto per \mathcal{F} .

(1.5.10) Definizione. Si chiama *tribù degli eventi anteriori a T* , e si denota con \mathcal{F}_T , la tribù costituita da tutti gli elementi A di \mathcal{F}_∞ tale che la funzione T_A definita in (1.2.1) sia un tempo d'arresto, ossia per ogni numero reale positivo t , si abbia

$$\{T_A \leq t\} = A \cap \{T \leq t\} \in \mathcal{F}_t.$$

È del tutto evidente che, se A è un evento contenuto in $\{T < \infty\}$, allora esso appartiene certamente a \mathcal{F}_T non appena la funzione T_A è un tempo d'arresto: in

effetti, se questa condizione è verificata, l'appartenenza di A alla tribù \mathcal{F}_∞ è automaticamente assicurata dalla relazione

$$A = \{T_A < \infty\} = \bigcup_n \{T_A \leq n\}.$$

È anche chiaro che, se A è una parte qualunque di Ω , affinché la funzione T_A sia un tempo d'arresto occorre e basta che il suo grafico sia progressivo grazie alla proposizione (1.5.8). È chiaro, inoltre, che se t è un numero reale positivo e T è il tempo d'arresto che assume t come unico valore costante, si ha $\mathcal{F}_T = \mathcal{F}_t$. Inoltre:

(1.5.11) Proposizione. *Per ogni coppia S, T di tempi d'arresto, con $S \leq T$, si ha $\mathcal{F}_S \subset \mathcal{F}_T$.*

Dimostrazione. Sia A un elemento di \mathcal{F}_S , ossia un elemento della tribù \mathcal{F}_∞ tale che la funzione S_A sia un tempo d'arresto. La relazione $T_A = S_A \vee T$ mostra allora che anche la funzione T_A è un tempo d'arresto. Ne segue che A è un elemento di \mathcal{F}_T .

(1.5.12) Proposizione. *Per ogni coppia S, T di tempi d'arresto, si ha*

$$\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T.$$

Dimostrazione. In effetti, dalla proposizione precedente segue che $\mathcal{F}_{S \wedge T}$ è contenuta dentro ciascuna delle due tribù \mathcal{F}_S e \mathcal{F}_T . Di conseguenza essa è contenuta nella loro intersezione. Per provare l'altra inclusione, si consideri un elemento A di $\mathcal{F}_S \cap \mathcal{F}_T$, cioè un elemento A di \mathcal{F}_∞ tale che ciascuna delle due funzioni S_A e T_A sia un tempo d'arresto. La relazione

$$(S \wedge T)_A = S_A \wedge T_A$$

mostra allora che la funzione $(S \wedge T)_A$ è anch'essa un tempo d'arresto: di conseguenza A appartiene alla tribù $\mathcal{F}_{S \wedge T}$.

(1.5.13) Proposizione. *Per ogni coppia S, T di tempi d'arresto, gli insiemi*

$$\{S < T\}, \quad \{S > T\}, \quad \{S = T\}$$

appartengono a $\mathcal{F}_{T \wedge S}$ e quindi appartengono a ciascuna delle due tribù $\mathcal{F}_S, \mathcal{F}_T$.

Dimostrazione. Occupiamoci dell'insieme $A = \{S < T\}$: le altre due dimostrazioni sono identiche. Poiché quest'insieme è contenuto in $\{S \wedge T < \infty\}$, è sufficiente verificare che la funzione $(S \wedge T)_A$ è un tempo d'arresto, ossia che il suo grafico è progressivo. Questo segue infatti dall'uguaglianza seguente:

$$\llbracket (S \wedge T)_A \rrbracket = \llbracket S \wedge T \rrbracket \cap \llbracket T \rrbracket^c.$$

(1.5.14) Corollario. *Ogni tempo d'arresto T è misurabile rispetto alla tribù \mathcal{F}_T .*

Dimostrazione. È sufficiente ricordare che, per ogni numero reale positivo t , la relazione $\{T < t\} \in \mathcal{F}_T$ segue dalla proposizione precedente.

Sia data una funzione T da Ω a valori in $[0, \infty]$. Si consideri l'applicazione biiettiva τ dell'insieme $\{T < \infty\}$ sul grafico di T , definita come

$$(1.5.15) \quad \tau(\omega) = (T(\omega), \omega).$$

Per ogni parte A di $\{T < \infty\}$ si ha chiaramente

$$\tau(A) = \llbracket T_A \rrbracket.$$

Supponiamo, inoltre, che T sia un tempo d'arresto. È allora chiaro che, affinché una parte A di $\{T < \infty\}$ appartenga a \mathcal{F}_T è necessario e sufficiente che $\llbracket T_A \rrbracket$ sia progressivo. In altri termini:

(1.5.16) Proposizione. *Sia T un tempo d'arresto, la biiezione τ di $\{T < \infty\}$ su $\llbracket T \rrbracket$ definita da (1.5.15) è bimisurabile considerando su $\{T < \infty\}$ la traccia di \mathcal{F}_T e su $\llbracket T \rrbracket$ la traccia della tribù progressiva.*

Siano dati un processo X e una funzione T da Ω a valori in $[0, \infty]$: si denota con X_T la funzione definita sull'insieme $\{T < \infty\}$ dalla relazione

$$(1.5.17) \quad X_T(\omega) = X(T(\omega), \omega).$$

La funzione X_T è dunque la composizione della biiezione τ da $\{T < \infty\}$ su $\llbracket T \rrbracket$ (definita da (1.5.15)) e la restrizione di X a $\llbracket T \rrbracket$.

(1.5.18) Proposizione. *La funzione X_T definita da (1.5.17) sull'insieme $\{T < \infty\}$ è misurabile rispetto alla traccia di \mathcal{F}_T su quest'insieme in ciascuno dei casi seguenti:*

- (a) *quando X è un processo progressivo e T un tempo d'arresto;*
- (b) *quando X è un processo adattato e T è un tempo d'arresto discreto.*

Dimostrazione. Nel caso (a), la conclusione segue immediatamente dalla proposizione precedente. Nel caso (b), si ponga $D = T(\Omega) \cup \{0\}$ e, per ciascun numero reale positivo t , si denoti con $\sigma(t)$ il più grande elemento dell'insieme $D \cap [0, t]$. Il processo Y definito come

$$Y(t, \omega) = X(\sigma(t), \omega)$$

è allora un processo adattato, continuo a destra e quindi progressivo. Esso coincide, inoltre, con X sul grafico di T e quindi si ha $X_T = Y_T$. Il caso (b) è così ricondotto al caso (a).

Applicando il risultato appena provato alla filtrazione deterministica si ottiene il corollario seguente:

(1.5.19) Corollario. *La funzione X_T è misurabile rispetto alla traccia di \mathcal{A} su $\{T < \infty\}$ in ciascuno dei casi seguenti:*

- (a) *quando X è un processo misurabile e T è una variabile aleatoria positiva;*
- (b) *quando X è un arbitrario processo e T è una variabile aleatoria positiva e discreta.*

Sullo spazio probabilizzato (Ω, \mathcal{A}, P) , sia X un processo reale e T una variabile aleatoria a valori in $[0, \infty]$. Si denota con $X^{|T}$ l'applicazione da $\mathbb{R}_+ \times \Omega$ a \mathbb{R} definita dalla relazione

$$X^{|T}(t, \omega) = X(t \wedge T(\omega), \omega).$$

Si dice, inoltre, che tale applicazione è ottenuta *arrestando* il processo X *all'istante* T .

(1.5.20) Proposizione. (a) *Se X è un processo adattato e T è un tempo d'arresto discreto, allora $X^{|T}$ risulta un processo adattato.*

(b) *Se X è un processo progressivo e T è un tempo d'arresto, allora $X^{|T}$ risulta un processo progressivo.*

(c) *Se X è un processo nullo in 0 e prevedibile e se T è un tempo d'arresto qualunque, allora $X^{|T}$ risulta un processo nullo in 0 e prevedibile.*

Dimostrazione. Cominciamo col provare che, in ciascuno dei tre casi, $X^{|T}$ è un processo adattato. Per fare questo è sufficiente ricordare che, per ogni numero reale positivo t , la funzione

$$(X^{|T})_t = X_{t \wedge T}$$

è misurabile rispetto alla tribù $\mathcal{F}_{t \wedge T}$ (si veda a questo proposito la proposizione (1.5.18)) e quindi anche rispetto alla tribù \mathcal{F}_t .

Per concludere basta scrivere $X^{|T}$ nella forma

$$X^{|T} = XI_{\llbracket 0, T \rrbracket} + (X_T)I_{\llbracket T, \infty \rrbracket}.$$

Difatti, il primo termine di questa decomposizione è un processo progressivo nel caso (b) e prevedibile nel caso (c) (potendolo scrivere in quest'ultimo caso nella forma $XI_{\llbracket 0, T \rrbracket}$). Il secondo termine è in entrambi i casi un processo prevedibile e nullo in 0 (essendo un processo adattato la cui restrizione a $\llbracket 0, \infty \rrbracket$ è un processo continuo a sinistra).

Applicando il risultato appena dimostrato alla filtrazione deterministica si ottiene immediatamente il corollario seguente:

(1.5.21) Corollario. *Se X è un processo misurabile e T è una variabile aleatoria positiva, $X^{|T}$ è un processo misurabile.*

II

Martingale

2.1 La nozione di martingala

Sia fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, una filtrazione \mathcal{F} avente come insieme dei tempi un arbitrario insieme ordinato I .

(2.1.1) Definizione. Un processo reale X su (Ω, \mathcal{A}, P) , avente come insieme dei tempi una parte J di I , è detto una *martingala* (relativa a \mathcal{F}) se valgono le condizioni seguenti:

- (a) X è adattato a \mathcal{F} ;
- (b) per ciascun elemento t di J , la variabile aleatoria X_t è integrabile;
- (c) per ogni coppia s, t di elementi di J , con $s \leq t$, la variabile aleatoria X_s è una versione della speranza condizionale di X_t rispetto a \mathcal{F}_s , cosa che si esprime tramite la relazione

$$X_s \in P[X_t | \mathcal{F}_s],$$

o, più impropriamente, se si conviene di denotare con X_s la classe in $L^1(P)$ avente X_s come rappresentante, $X_s = P[X_t | \mathcal{F}_s]$.

Attraverso le proprietà elementari della speranza condizionale è immediato verificare che la condizione (c) della definizione (2.1.1) è equivalente alla condizione:

- (c') Per ogni coppia s, t di elementi di I , con $s \leq t$, si ha

$$P[X_t - X_s | \mathcal{F}_s] = 0.$$

Quest'ultima condizione suggerisce la seguente definizione:

(2.1.2) Definizione. Un processo reale X su (Ω, \mathcal{A}, P) , avente come insieme dei tempi una parte J di I , è detto una *sottomartingala* (risp. *sopramartingala*) se valgono le condizioni (a) e (b) della definizione (2.1.1) e se vale inoltre la condizione:

(c'') Per ogni coppia s, t di elementi di J , con $s \leq t$, si ha:

$$P[X_t - X_s | \mathcal{F}_s] \geq 0 \quad (\text{risp. } P[X_t - X_s | \mathcal{F}_s] \leq 0).$$

È chiaro che un processo reale X su (Ω, \mathcal{A}, P) avente come insieme dei tempi una parte J di I è una sopramartingala se, e soltanto se, il processo $-X$ è una sottomartingala. È altrettanto chiaro che il processo X è una martingala se, e soltanto se, esso risulta contemporaneamente una sottomartingala e una sopramartingala. Inoltre, è facile riconoscere che, affinché X sia una martingala (risp. una sottomartingala) è necessario che la funzione $t \mapsto P[X_t]$ sia costante (risp. crescente).

Siano X, Y due martingale (rispetto alla medesima filtrazione \mathcal{F}). Si riconosce allora facilmente che anche $aX + bY$ è una martingala per ogni coppia a, b di numeri reali. Dunque, l'insieme costituito da tutte le martingale su (Ω, \mathcal{A}, P) (rispetto a \mathcal{F}) forma uno spazio vettoriale. Un risultato analogo vale per le sottomartingale, per ogni coppia a, b di numeri reali positivi. Inoltre, se X, Y sono sottomartingale, tale è $X \vee Y$. Infine, se X è una martingala (risp. sottomartingala) e f è una funzione convessa (risp. convessa e crescente), dalla disuguaglianza di Jensen condizionale segue che $f(X)$ è una sottomartingala. Di conseguenza, se X è una martingala, ciascuno dei processi $|X|, X^+, X^-$ risulta essere una sottomartingala; se X è una sottomartingala, tale è X^+ .

L'esempio che segue offre un metodo molto comodo per costruire martingale a partire da una variabile aleatoria integrabile fissata.

(2.1.3) Esempio. Sia Z una variabile aleatoria reale ed integrabile sullo spazio probabilizzato (Ω, \mathcal{A}, P) e sia \mathcal{F} una filtrazione su di esso, avente un insieme ordinato I come insieme dei tempi. Per ciascun indice t , si denoti con X_t una versione di $P[Z | \mathcal{F}_t]$. Allora il processo X è chiaramente una martingala rispetto all'assegnata filtrazione \mathcal{F} .

Questo esempio suggerisce anche la seguente definizione.

(2.1.4) Definizione. Sia X una martingala (risp. sottomartingala) su (Ω, \mathcal{A}, P) avente un arbitrario insieme I come insieme dei tempi. Si supponga che I non contenga un elemento massimale. Si dice allora che X è *chiusa* da una variabile aleatoria Z se essa è integrabile e se, per ciascun indice t , è verificata la relazione

$$X_t \in P[Z | \mathcal{F}_t] \quad (\text{risp. } X_t \leq P[Z | \mathcal{F}_t]).$$

In tal caso si dice anche che Z *chiude* la martingala X .

La nozione di martingala chiusa può essere così interpretata: aggiungiamo a I un elemento più grande di ogni suo elemento (che denoteremo con ∞), e si ponga

$\bar{I} = I \cup \{\infty\}$. Si denoti con \mathcal{F}_∞ la minima tribù contenente ciascuna delle \mathcal{F}_t , e tale che Z sia misurabile rispetto ad essa. Allora la relazione $X_t = P[Z | \mathcal{F}_t]$ significa che il processo $(X_t)_{t \in \bar{I}}$ ottenuto ponendo $X_\infty = Z$ è una martingala.

2.2 Martingale a tempi discreti

Nel resto di questa sezione supporremo sempre che, sullo spazio probabilizzato (Ω, \mathcal{A}, P) , sia fissata una filtrazione $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$. Diremo che un processo è *discreto* se esso ha come insieme dei tempi una parte numerabile di \mathbb{R}_+ .

Sotto queste ipotesi, un processo discreto $(V_n)_{n \geq 0}$ si dice *prevedibile* se, per ciascun intero positivo n , V_n è misurabile rispetto alla tribù \mathcal{F}_{n-1} . Si tratta evidentemente della nozione corrispondente, nel caso discreto, alla nozione di processo prevedibile già incontrata nel capitolo precedente. Se X è un processo discreto e adattato, e se V è prevedibile, si denota con $V \cdot X$ il processo così definito

$$(V \cdot X)_n = X_0 + V_1(X_1 - X_0) + \cdots + V_n(X_n - X_{n-1}).$$

Chiameremo questo processo *trasformata di Burkholder di X rispetto a V* ; in effetti si tratta di una forma molto semplice della nozione di integrale stocastico che studieremo nei capitoli successivi.

(2.2.1) Proposizione. *Sia X una martingala (risp. sottomartingala) discreta, e sia V un processo prevedibile (risp. prevedibile e positivo) discreto. Se, per ciascun intero positivo n , $(V \cdot X)_n$ è integrabile, allora $V \cdot X$ è anch'essa una martingala (risp. sottomartingala).*

Dimostrazione. Nel caso in cui X sia una sottomartingala, basta verificare che, per ciascun intero positivo n , si ha

$$P[(V \cdot X)_{n+1} - (V \cdot X)_n | \mathcal{F}_n] \geq 0.$$

In effetti, si ha

$$\begin{aligned} P[(V \cdot X)_{n+1} - (V \cdot X)_n | \mathcal{F}_n] &= P[V_{n+1}(X_{n+1} - X_n) | \mathcal{F}_n] \\ &= V_{n+1}P[X_{n+1} - X_n | \mathcal{F}_n] \geq 0. \end{aligned}$$

Il caso in cui X sia una martingala è analogo.

Osserviamo che la definizione precedente ha senso anche quando V sia un processo discreto avente un arbitrario insieme discreto I (non necessariamente eguale a \mathbb{N}) come insieme dei tempi.

Ricordiamo che si chiama *tempo d'arresto discreto* un tempo d'arresto T a valori in un insieme discreto I . Nel caso particolare in cui il processo V sia definito da

$$V_i = I_{\{i \leq T\}}, \quad \text{per ogni } i \in I$$

con T tempo d'arresto discreto, è facile riconoscere che il processo $V \cdot X$ coincide con $X^{[T]}$ cosicché anche quest'ultimo processo risulta essere una martingala. Abbiamo così provato il corollario seguente:

(2.2.2) Corollario. Sia X una sottomartingala (risp. martingala) discreta, e sia T un tempo d'arresto discreto. Allora $X^{|T|}$ è anch'essa una sottomartingala (risp. martingala) discreta.

(2.2.3) Teorema (d'arresto discreto). Sia assegnata una sottomartingala discreta X (avente I come insieme dei tempi) e siano S, T una coppia di tempi d'arresto discreti e limitati (a valori in I), con $S \leq T$. Allora X_S e X_T sono integrabili e si ha

$$X_S \leq P[X_T | \mathcal{F}_S].$$

Nel caso particolare in cui X sia per giunta una martingala, la relazione precedente diviene un'uguaglianza.

Dimostrazione. Per semplicità supporremo che sia $I = \mathbb{N}$. Sia k un intero maggiore o eguale a T . Allora $|X_S|$ e $|X_T|$ sono maggiorate da $|X_0| + \dots + |X_k|$, per cui X_S e X_T sono integrabili. Sia adesso A un elemento della tribù \mathcal{F}_S ; per ogni intero positivo j inferiore o eguale a k , l'insieme $A \cap \{S = j\}$ appartiene a \mathcal{F}_j , e dunque

$$\int_{A \cap \{S=j\}} (X_k - X_S) dP = \int_{A \cap \{S=j\}} (X_k - X_j) dP \geq 0.$$

Sommando per ciascun indice j , per arbitrarietà dell'insieme A appartenente a \mathcal{F}_S , si ottiene la relazione $P[X_k - X_S | \mathcal{F}_S] \geq 0$. Per concludere basta applicare questo risultato al processo arrestato $X^{|T|}$.

(2.2.4) Lemma massimale discreto. Siano assegnate una sottomartingala discreta X (avente I come insieme dei tempi), un elemento t in I , una parte finita D di $[0, t] \cap I$, e la variabile aleatoria U definita da

$$U = \sup_{r \in D} X_r.$$

Si ha allora, per ciascun numero reale positivo x ,

$$xP\{U \geq x\} \leq \int_{\{U > x\}} X_t dP.$$

Dimostrazione. Fissato un numero reale positivo x , si ponga, per ciascun elemento ω di Ω ,

$$S(\omega) = \inf\{r \in D : X_r(\omega) > x\}.$$

Si ha allora, per ogni numero reale positivo s ,

$$\{S \leq s\} = \bigcup_{r \in D, r \leq s} \{X_r > x\}.$$

Questo basta per concludere che S è un tempo d'arresto (a valori in $D \cup \{\infty\}$). La relazione precedente fornisce inoltre, per $s = t$,

$$\{S \leq t\} = \bigcup_{r \in D} \{X_r > x\},$$

cioè

$$\{S < \infty\} = \{U > x\}.$$

Ne segue, dal fatto che X è una sottomartingala,

$$\begin{aligned} xP\{U > x\} &= xP\{S < \infty\} \\ &\leq \int_{\{S < \infty\}} X_S \, dP \\ &= \sum_{r \in D} \int_{\{S=r\}} X_r \, dP \\ &\leq \sum_{r \in D} \int_{\{S=r\}} X_t \, dP = \int_{\{S < \infty\}} X_t \, dP, \end{aligned}$$

e ciò basta per concludere.

(2.2.5) Teorema (disuguaglianza di Doob sul numero di discese). *Siano X una sottomartingala discreta (avente I come insieme dei tempi), t un numero reale positivo in I , D una parte finita di $[0, t] \cap I$, a, b una coppia di numeri reali, con $a < b$. Si ponga, per ogni ω ,*

$$\begin{aligned} S_1(\omega) &= \inf\{s \in D : X_s(\omega) \geq b\}, \\ T_1(\omega) &= \inf\{s \in D : s > S_1(\omega), X_s(\omega) \leq a\}, \\ S_k(\omega) &= \inf\{s \in D : s > T_{k-1}(\omega), X_s(\omega) \geq b\}, \\ T_k(\omega) &= \inf\{s \in D : s > S_{k-1}(\omega), X_s(\omega) \leq a\}. \end{aligned}$$

Le funzioni S_k e T_k sono allora tempi d'arresto e, se si pone

$$J = J(a, b, D) = \sum_{k \geq 1} I_{\{T_k < \infty\}},$$

si ha

$$P[J] \leq (b - a)^{-1} P[(X_t - b)^+].$$

Dimostrazione. Innanzitutto si riconosce facilmente, ragionando per induzione, che S_k e T_k sono tempi d'arresto, con

$$S_k \leq T_k \leq S_{k+1} \leq T_{k+1}.$$

Per ogni coppia di tempi d'arresto discreti S, T , con $S \leq T$, il teorema d'arresto discreto (2.2.3) mostra che, se A è un elemento di \mathcal{F}_S e B è una parte \mathcal{F}_S -misurabile di A , allora

$$0 \leq \int_A (X_T - X_S) \, dP = \int_{A \setminus B} (X_T - X_S) \, dP + \int_B (X_T - X_S) \, dP,$$

da cui segue

$$\int_B (X_S - X_T) \, dP \leq \int_{A \setminus B} (X_T - X_S) \, dP.$$

Per ogni fissato intero positivo k , è facile riconoscere che

$$(2.2.6) \quad \int_{\{T_k < \infty\}} (X_{S_k} - X_{T_k}) dP \geq (b - a)P\{T_k < \infty\}.$$

Applicando la relazione (2.2.6) alla coppia di tempi d'arresto S_k e $T_k \wedge t$ e ponendo $A = \{S_k < \infty\}$ e $B = \{T_k < \infty\}$, si ha

$$(b - a)P\{T_k < \infty\} \leq \int_{\{S_k < \infty, T_k = \infty\}} (X_t - X_{S_k}) dP \leq \int_{\{S_k < \infty, T_k = \infty\}} (X_t - b) dP.$$

Sommando la relazione precedente, al variare di k , e denotato con C la riunione degli insiemi $\{S_k < \infty, T_k = \infty\}$ (che sono a due a due disgiunti), si ottiene:

$$(b - a)P[J] \leq \int_C (X_t - b) dP \leq P[(X_t - b)^+].$$

2.3 Martingale a tempi continui

D'ora in poi, sullo spazio probabilizzato (Ω, \mathcal{A}, P) , sarà sempre fissata una filtrazione \mathcal{F} avente \mathbb{R}_+ come insieme dei tempi. Inoltre, salvo avviso contrario, ogni processo avrà \mathbb{R}_+ come insieme dei tempi. Osserviamo che, se D è una parte numerabile di \mathbb{R}_+ , e $(X_t)_{t \geq 0}$ è una sottomartingala, è facile riconoscere che $(X_t)_{t \in D}$ è una sottomartingala discreta, avente D come insieme dei tempi.

(2.3.1) Lemma. *Sia X una sottomartingala e sia (S_n) una successione decrescente di tempi d'arresto discreti e limitati. La successione (X_{S_n}) è allora uniformemente integrabile.*

Dimostrazione. Cominciamo con l'osservare che le successioni $(P[X_{S_n}])$ e $(P[X_{S_n}^+])$ sono decrescenti e minorate rispettivamente dai numeri $P[X_0]$ e $P[X_0^+]$. Pertanto anche la successione $(P[X_{S_n}^-])$ è limitata (essendo differenza di due successioni limitate).

Per ogni numero reale positivo ε , è possibile trovare un intero positivo k tale che, per ogni intero positivo n , con $n \geq k$, si abbia

$$0 \leq P[X_{S_k}] - P[X_{S_n}] \leq \varepsilon.$$

Sia ora c una costante reale positiva tale che, per ogni n con $n < k$, si abbia

$$(2.3.2) \quad \int_{\{|X_{S_n}| > c\}} |X_{S_n}| dP < \varepsilon.$$

Poiché X è una sottomartingala, denotato con D_n l'insieme (numerabile) dei valori di S_n , e denotato con D la riunione dei D_n , è possibile applicare il teorema d'arresto

discreto (2.2.3) alla sottomartingala discreta (X_{S_n}) e, osservando che la tribù \mathcal{F}_{S_n} risulta contenuta in \mathcal{F}_{S_k} per ciascun intero positivo n con $n \geq k$, si ha

$$\begin{aligned} \int_{\{|X_{S_n}|>c\}} |X_{S_n}| dP &= \int_{\{X_{S_n}>c\}} X_{S_n} dP - \int_{\{X_{S_n}<-c\}} X_{S_n} dP \\ &= \int_{\{X_{S_n}>c\}} X_{S_n} dP + \int_{\{X_{S_n}\geq -c\}} X_{S_n} dP - P[X_{S_n}] \\ &\leq \int_{\{X_{S_n}>c\}} X_{S_k} dP + \int_{\{X_{S_n}\geq -c\}} X_{S_k} dP - P[X_{S_n}] \\ &\leq \int_{\{X_{S_n}>c\}} X_{S_k} dP + \int_{\{X_{S_n}\geq -c\}} X_{S_k} dP - P[X_{S_k}] + \varepsilon \\ &= \int_{\{|X_{S_n}|>c\}} |X_{S_k}| dP + \varepsilon. \end{aligned}$$

Inoltre, per la disuguaglianza di Markov, per ciascun intero n , si ha

$$P\{|X_{S_n}| > c\} \leq c^{-1} \sup_n P[|X_{S_n}|]$$

e l'ultimo termine converge verso 0 al tendere di c all'infinito perché la successione $(P[|X_{S_n}|])$ è limitata. Dunque, mettendo insieme quest'osservazione con (2.3.2), per c sufficientemente grande, e per ciascun intero positivo n , è verificata la relazione

$$\int_{\{|X_{S_n}|>c\}} |X_{S_n}| dP \leq 2\varepsilon.$$

Questo completa la dimostrazione.

(2.3.3) Teorema (d'arresto). *Sia X una sottomartingala continua a destra. Sia S, T una coppia di tempi d'arresto limitati, con $S \leq T$. Allora X_S e X_T sono integrabili e si ha*

$$X_S \leq P[X_T | \mathcal{F}_S].$$

Nel caso particolare in cui X sia per giunta una martingala, la relazione precedente diviene un'uguaglianza.

Dimostrazione. Sia D_n l'insieme dei numeri diadici di stadio n -esimo (ossia dei numeri della forma $k2^{-n}$, con k intero positivo). Si ponga, per ciascun intero positivo n ,

$$S_n(\omega) = \inf\{t \in D_n : S(\omega) \leq t\}.$$

Chiaramente, per ciascun intero positivo n , S_n è un tempo d'arresto rispetto alla filtrazione $(\mathcal{F}_t)_{t \in D_n}$. Inoltre, la successione (S_n) è decrescente e ammette S come inviluppo inferiore. Di conseguenza (X_{S_n}) è una successione uniformemente integrabile (si veda (2.3.1)). Poiché essa converge puntualmente verso X_S , si vede che X_S è integrabile, e che la convergenza di X_{S_n} verso X_S ha luogo in \mathcal{L}^1 . Si ragiona in modo analogo per costruire una successione decrescente (T_n) di tempi d'arresto, con $S_n \leq T_n$, aventi T come inviluppo inferiore, in modo tale che X_{T_n} converga in \mathcal{L}^1 verso X_T . Sia poi A un elemento della tribù \mathcal{F}_S . Siccome S è maggiorata da ciascuna delle S_n , A appartiene a ciascuna delle tribù \mathcal{F}_{S_n} . Applicando dunque il teorema d'arresto discreto (2.2.3) si ottiene

$$\int_A X_{S_n} dP \leq \int_A X_{T_n} dP.$$

Per concludere basta passare al limite ed utilizzare la convergenza in \mathcal{L}^1 .

(2.3.4) Corollario. *Sia X una martingala e sia T un tempo d'arresto. Allora anche il processo arrestato $X|_T$ è una martingala.*

Dimostrazione. Sia s, t una coppia di numeri reali positivi, con $s \leq t$. Per concludere basta osservare che, per il teorema d'arresto (2.3.3), si ha

$$P[X_{t \wedge T} - X_{s \wedge T} | \mathcal{F}_s] = 0.$$

(2.3.5) Lemma massimale. *Sia X una martingala continua a destra, chiusa dalla variabile aleatoria Z . Si ponga*

$$X_\infty^* = \sup_{t \in \mathbb{R}_+} |X_t|.$$

Si ha allora, per ciascun numero reale e positivo x , e ciascun numero reale strettamente positivo λ ,

$$x^\lambda P\{X_\infty^* > x\} \leq \int_{\{X_\infty^* > x\}} |Z|^\lambda dP.$$

Dimostrazione. Sia D una parte numerabile e densa di \mathbb{R}_+ , e sia (D_n) una successione crescente di parti finite di D , avente D come riunione. Si ponga, per ogni n ,

$$U_n = \sup_{s \in D_n} |X_s|,$$

e denotiamo con t_n il più grande elemento di D_n . Applicando il lemma massimale discreto (2.2.4) alla sottomartingala $|X|^\lambda$, si trova

$$(2.3.6) \quad x^\lambda P\{U_n > x\} \leq \int_{\{U_n > x\}} |X_{t_n}|^\lambda dP \leq \int_{\{U_n > x\}} |Z|^\lambda dP.$$

D'altra parte si ha

$$\sup_n U_n = \sup_{t \in D} |X_t| = X_\infty^*,$$

e la successione (U_n) è crescente. La conclusione si ottiene dunque passando al limite in (2.3.6).

Allo scopo di caratterizzare le martingale chiuse mediante l'uniforme integrabilità, occorre premettere il seguente lemma relativo alla regolarità delle traiettorie di una martingala continua a destra.

(2.3.7) Lemma. *Sia X una martingala continua a destra. Allora quasi ogni traiettoria di X è regolare. Supponiamo inoltre che X sia limitata in L^1 , cioè tale che si abbia*

$$\sup_t P[|X_t|] < \infty.$$

Allora quasi ogni traiettoria di X è limitata e dotata di limite all'infinito.

Dimostrazione. La decomposizione $X = X^+ - X^-$ permette di ricondurci al caso in cui X sia una martingala positiva. Inoltre, per ogni fissato ω in Ω , se la restrizione di $X(\cdot, \omega)$ a \mathbb{Q}_+ è regolare (risp. limitata e dotata di limite all'infinito), tale è anche l'intera traiettoria $X(\cdot, \omega)$. Basta dunque provare che, per quasi ogni ω in Ω , la restrizione di $X(\cdot, \omega)$ a una parte numerabile e non limitata D di \mathbb{R}_+ è regolare (risp. limitata e dotata di limite all'infinito, nel caso particolare in cui X sia limitata in L^1).

A questo scopo sia (D_n) una successione crescente di parti finite di D , la cui riunione coincida con D , e denotiamo con t_n il più grande elemento di D_n . Si ponga inoltre

$$U_n = \sup_{s \in D_n} X_s.$$

La successione (U_n) è allora crescente, e il suo inviluppo superiore è la variabile aleatoria

$$U = \sup_{s \in D} X_s.$$

Proviamo che questa variabile aleatoria è quasi certamente finita. Il lemma massimale (2.2.4) fornisce, per ogni intero positivo n , e ogni numero reale $x > 0$,

$$P\{U_n > x\} \leq \frac{1}{x} \int X_{t_n} dP \leq \frac{c}{x}.$$

Ne segue $P\{U_n > x\} \leq c/x$, da cui $P\{U = \infty\} = 0$. Poniamo dunque (con le notazioni utilizzate in (2.2.5))

$$J_n(a, b) = J(a, b, D_n),$$

e denotiamo con $J(a, b)$ l'inviluppo superiore della successione crescente $(J_n(a, b))$. Si ha allora, per il teorema (2.2.5),

$$P[J_n(a, b)] \leq (b - a)^{-1} P[(X_{t_n} - b)^+],$$

dunque anche

$$P[J(a, b)] \leq (b - a)^{-1} \sup_{s \in D} P[(X_s - b)^+] < \infty.$$

Di conseguenza l'insieme

$$A = \bigcup_{a, b \in \mathbb{Q}_+, a < b} \{J(a, b) = \infty\}$$

è trascurabile.

Sia ω un elemento di Ω , tale che la restrizione della traiettoria $X(\cdot, \omega)$ all'insieme D sia limitata, ma non regolare a sinistra (oppure priva di limite all'infinito). Esiste allora un elemento t di $]0, \infty]$ verificante la relazione

$$\liminf_{s \uparrow t, s \in D} X_s(\omega) < \limsup_{s \uparrow t, s \in D} X_s(\omega).$$

Si può dunque scegliere una coppia a, b di numeri reali positivi, con $a < b$, e una successione strettamente crescente (s_k) di elementi di D , con $\lim_k s_k = t$, dimodoché si abbia, per ogni intero positivo h ,

$$X_{s_{2h}}(\omega) \geq b, \quad X_{s_{2h+1}}(\omega) \leq a.$$

Per ogni intero h , esiste un intero n_h tale che i numeri s_1, \dots, s_{2h} appartengano tutti a D_{n_h} . Si ha allora $J_{n_h}(a, b)(\omega) \geq h$ e, di conseguenza, $J(a, b)(\omega) = \infty$. Questo prova che ω appartiene all'insieme trascurabile A . Allo stesso modo si dimostra che A contiene l'insieme degli elementi ω di Ω tali che la traiettoria $X(\cdot, \omega)$ sia limitata, ma non regolare a destra. Il lemma è così dimostrato.

Veniamo ora all'attesa caratterizzazione delle martingale chiuse.

(2.3.8) Proposizione. *Sia X una martingala continua a destra. Le condizioni che seguono sono allora equivalenti:*

- (a) X è chiusa.
- (b) La famiglia $(X_t)_{t \in \mathbb{R}_+}$ è uniformemente integrabile.

Inoltre, se le due condizioni sono verificate, X_t converge quasi certamente e in \mathcal{L}^1 (quando t tende verso l'infinito) verso una variabile aleatoria reale e misurabile rispetto alla tribù \mathcal{F}_∞ che la chiude.

Dimostrazione. (a) \Rightarrow (b): Supponiamo che la martingala X sia chiusa dalla variabile aleatoria Z . Per ogni t , X_t è allora una versione della speranza condizionale $P[Z | \mathcal{F}_t]$. Grazie all'uguaglianza $P[Z | \mathcal{F}_t] = P[Z^+ | \mathcal{F}_t] - P[Z^- | \mathcal{F}_t]$, possiamo ricondurci al caso in cui le variabili aleatorie Z e X_t sono tutte positive. Fissiamo allora un numero reale e positivo ε , e scegliamo un numero reale positivo δ tale che, per ogni elemento A di \mathcal{A} , con $P(A) < \delta$, risulti $\int_A Z dP < \varepsilon$. Scegliamo, poi, un numero reale e positivo c tale che si abbia $c^{-1}P[Z] < \delta$. Per la disuguaglianza di Markov, si ha allora, per ogni numero reale e positivo t ,

$$P\{X_t > c\} \leq c^{-1}P[X_t] = c^{-1}P[Z] < \delta,$$

e quindi

$$\int_{\{X_t > c\}} X_t dP = \int_{\{X_t > c\}} Z dP < \varepsilon.$$

Questo prova che la famiglia (X_t) è uniformemente integrabile.

(b) \Rightarrow (a): Si supponga che la famiglia (X_t) sia uniformemente integrabile, e dunque limitata in \mathcal{L}^1 . Ne segue allora (si veda (2.3.7)) che quasi ogni traiettoria $X(\cdot, \omega)$ è limitata e dotata di limite all'infinito. Di conseguenza, se si denota

con V la variabile aleatoria coincidente con la variabile aleatoria (misurabile rispetto a \mathcal{F}_∞)

$$V = \limsup_{t \rightarrow \infty, t \in \mathbb{Q}_+} X_t$$

laddove questo limite è finito, e con 0 altrove, si vede facilmente che V è misurabile rispetto a \mathcal{F}_∞ , e che X_t converge verso V (quando t tende verso l'infinito), non solo quasi certamente, ma anche nel senso della convergenza in \mathcal{L}^1 . Resta soltanto da provare che V chiude la martingala X . Consideriamo a questo scopo un numero reale positivo s e un elemento A di \mathcal{F}_s . Si ha allora

$$\int_A X_s dP = \int_A X_{s+t} dP,$$

dalla quale si deduce (grazie alla convergenza di X_t in \mathcal{L}^1)

$$\int_A X_s dP = \int_A V dP.$$

La proposizione è così dimostrata.

Data una variabile aleatoria integrabile Z , ci domandiamo se esiste una martingala che sia chiusa da Z e che sia per giunta continua a destra. Quanto la filtrazione verifica le condizioni abituali la risposta è affermativa, come mostra la proposizione seguente:

(2.3.9) Proposizione. *Supponiamo che la filtrazione \mathcal{F} verifichi le condizioni abituali. Esiste allora, per ogni variabile aleatoria Z integrabile, una martingala X chiusa da Z e continua a destra. Inoltre, si può fare in modo che X sia a traiettorie limitate, regolari e dotate di limite all'infinito (e che X sia limitata se tale è Z).*

Dimostrazione. La decomposizione $Z = Z^+ - Z^-$ permette di ricondurci al caso in cui Z sia positiva.

Cominciamo con il costruire una martingala Y chiusa rispetto a Z : è sufficiente per questo scegliere, per ciascun t , una versione Y di $P[Z | \mathcal{F}_t]$ che sia ovunque positiva. Consideriamo poi una parte numerabile D di \mathbb{R}_+ , densa in \mathbb{R}_+ . Allo stesso modo della dimostrazione di (2.3.7), si prova che, quasi ogni traiettoria di Y ha una restrizione a D che è limitata, regolare e dotata di limite all'infinito. A meno di modificare Y_t , con $t \in D$, su di un insieme trascurabile (appartenente a \mathcal{F}_0 per ipotesi), si potrà supporre che la proprietà precedente sia verificata per tutte le traiettorie di Y . Si ponga allora, per ogni elemento (s, ω) di $\mathbb{R}_+ \times \Omega$:

$$X(s, \omega) = \lim_{t \downarrow s, t \in D} Y(t, \omega).$$

Si definisce così un processo X , le cui traiettorie sono limitate, continue a destra, limitate, e dotate di limite all'infinito.

È chiaro che, per ogni s , la variabile aleatoria X_s è misurabile rispetto alla tribù \mathcal{F}_s^+ , che, per ipotesi, coincide con \mathcal{F}_s . Il processo X è dunque adattato. Resta

da provare che X è una martingala. A questo scopo fissiamo un numero reale e positivo s e un elemento A di \mathcal{F}_s . Si ha allora, per ogni numero reale t , con $t \geq s$,

$$\int_A Y_t dP = \int_A Z dP.$$

Inoltre, la famiglia (Y_t) è uniformemente integrabile (si veda (2.3.8)), e dunque, per definizione, si ha

$$\lim_{t \downarrow s, t \in D} Y_t = X_s.$$

Ne segue che

$$\int_A X_s dP = \int_A Z dP,$$

e ciò basta per concludere.

Osserviamo che, se U è una funzione numerica positiva e misurabile sullo spazio misurabile (Ω, \mathcal{A}) , e se μ è una misura positiva e σ -finita sullo stesso spazio, per ogni numero reale $p \geq 1$, si ha

$$(2.3.10) \quad \int U^p d\mu = \int_0^\infty px^{p-1} \mu\{U > x\} dx.$$

Per ottenere questa formula, basta scrivere U^p nella forma

$$\int_0^{U(\omega)} px^{p-1} dx$$

e applicare il teorema di Fubini:

$$\int U^p(\omega) \mu(d\omega) = \int \mu(d\omega) \int_0^{U(\omega)} px^{p-1} dx = \int px^{p-1} \mu\{\omega : U(\omega) > x\} dx.$$

La formula (2.3.10) permette di dimostrare facilmente il seguente lemma.

(2.3.11) Lemma. *Si consideri, sullo spazio probabilizzato (Ω, \mathcal{A}, P) , una coppia U, V di variabili aleatorie positive, con V quasi certamente finita, e si supponga che si abbia*

$$(2.3.12) \quad xP\{U > x\} \leq \int_{\{U > x\}} V dP$$

per quasi ogni numero reale positivo x . Si ha allora

$$(2.3.13) \quad \begin{aligned} P[U^2] &\leq 2P[UV], \\ P[U^2] &\leq 4P[V^2]. \end{aligned}$$

Dimostrazione. Si denoti con ν la misura $V \cdot P$. L'ipotesi (2.3.12) si traduce nella disuguaglianza

$$xP\{U > x\} \leq \nu\{U > x\}.$$

Integrando i due membri di questa relazione, si trova

$$\int_0^\infty xP\{U > x\} dx \leq \int_0^\infty \nu\{U > x\} dx.$$

Ora, la relazione (2.3.10) (ove si sia posto $p = 2$ e $\mu = P$) mostra che l'integrale a primo membro della relazione precedente è uguale a $\frac{1}{2}P[U^2]$. La stessa relazione (ove si sia posto $p = 1$ e $\mu = \nu$) mostra che l'integrale che figura a secondo membro di (2.3.13) è uguale a $\int U d\nu$, cioè a $\int UV dP$. La prima disuguaglianza è allora provata. Se adesso si applica, per ogni numero reale e positivo c , il risultato ottenuto alla variabile aleatoria $U \wedge c$ (che verifica ancora le ipotesi (2.3.12)), si trova

$$P[(U \wedge c)^2] \leq 2P[(U \wedge c)V] \leq 2\sqrt{P[(U \wedge c)^2]}\sqrt{P[V^2]},$$

dalla quale si deduce

$$P[(U \wedge c)^2] \leq 4P[V^2].$$

Non resta che fare tendere c verso l'infinito per ottenere (2.3.13).

Il lemma appena dimostrato permette di dedurre immediatamente dal lemma massimale (2.3.5) la *disuguaglianza di Doob*:

(2.3.14) Teorema (disuguaglianza di Doob). *Sia X una martingala continua a destra, chiusa dalla variabile aleatoria Z . Si ponga*

$$(2.3.15) \quad X_\infty^* = \sup_{t \in \mathbb{R}_+} |X_t|.$$

Si ha allora

$$P[(X_\infty^*)^2] \leq 4P[Z^2].$$

2.4 Martingale di quadrato integrabile

La disuguaglianza di Doob (2.3.14) permette di dimostrare una caratterizzazione importante della classe delle martingale continue a destra che sono chiuse da una variabile aleatoria di quadrato integrabile.

(2.4.1) Proposizione. *Sia X una martingala continua a destra. Le condizioni che seguono sono allora equivalenti:*

- (a) X è chiusa da una variabile aleatoria di quadrato integrabile;
- (b) La famiglia (X_t) è dominata in \mathcal{L}^2 (ossia la variabile aleatoria X_∞^* , definita da (2.3.15), è di quadrato integrabile);
- (c) La famiglia (X_t) è limitata in \mathcal{L}^2 (ossia $\sup_{t \in \mathbb{R}_+} P[X_t^2] < \infty$).

Inoltre, se queste condizioni sono verificate, (X_t) converge quasi certamente e in \mathcal{L}^2 (al tendere di t verso l'infinito) verso una variabile aleatoria V , misurabile rispetto a \mathcal{F}_∞ , e di quadrato integrabile, che chiude la martingala e che verifica la relazione

$$(2.4.2) \quad P[V^2] = \lim_{t \rightarrow \infty} P[X_t^2] = \sup_{t \in \mathbb{R}_+} P[X_t^2].$$

Dimostrazione. L'implicazione (a) \Rightarrow (b) segue da (2.3.14). L'implicazione (b) \Rightarrow (c) è evidente. Basta dunque dimostrare l'implicazione (c) \Rightarrow (a) e l'asserzione finale: a questo scopo supponiamo che la famiglia (X_t) sia limitata in \mathcal{L}^2 . Essa è allora uniformemente integrabile. Ne risulta dunque da (2.3.8) che, al tendere di t verso l'infinito, X_t converge quasi certamente verso una variabile aleatoria V , misurabile rispetto a \mathcal{F}_∞ che chiude la martingala. Inoltre, grazie al lemma di Fatou, si ha

$$P[V^2] \leq \liminf_{t \rightarrow \infty} P[X_t^2] < \infty,$$

per cui V è di quadrato integrabile, e l'asserzione (a) è provata. D'altronde, in virtù della condizione (b), la convergenza di X_t verso V ha luogo in \mathcal{L}^2 . Questo prova la relazione (2.4.2), dove l'ultima uguaglianza è dovuta al fatto che $t \mapsto P[X_t^2]$ è crescente, essendo X^2 una sottomartingala.

(2.4.3) Definizione. Si chiama *martingala di quadrato integrabile* ogni martingala X continua a destra e regolare, verificante le condizioni equivalenti (a), (b), (c) della proposizione precedente.

(2.4.4) Osservazione. Se supponiamo verificate le condizioni abituali, allora, per ogni variabile aleatoria Z di quadrato integrabile, esiste una martingala di quadrato integrabile X (unica a meno di processi indistinguibili) chiusa da Z (si veda (2.3.9)). Inoltre, questa martingala ammette come variabile aleatoria che la chiude una versione della speranza condizionale $P[Z | \mathcal{F}_\infty]$.

L'osservazione precedente si può riassumere nel risultato seguente:

(2.4.5) Proposizione. *Per ogni variabile aleatoria V , misurabile rispetto a \mathcal{F}_∞ e di quadrato integrabile, esiste un'unica martingala di quadrato integrabile (unica a meno di processi indistinguibili) chiusa da V .*

III

Il processo di Wiener

3.1 Variabili aleatorie gaussiane reali

(3.1.1) Definizione. Si chiama *legge normale ridotta* e si denota con $\mathcal{N}(0, 1)$ la legge sullo spazio misurabile $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ avente come densità la funzione

$$x \mapsto \frac{1}{\sqrt{2\pi}} \exp(-x^2/2).$$

La legge normale ridotta è simmetrica e dunque ha momento del prim'ordine nullo. Inoltre, essa ha momento del secondo ordine eguale a 1 (come si riconosce facilmente con una semplice integrazione per parti). Osservato che la funzione $x \mapsto x^k \exp(-x^2/2)$ è integrabile per ogni intero k , è facile provare che la legge normale ridotta ha momento di qualunque ordine finito. Si riconosce facilmente, inoltre, che la sua funzione caratteristica è la funzione così definita su \mathbb{R} :

$$\varphi(t) = \exp(-t^2/2).$$

(3.1.2) Definizione. Assegnata una coppia m, σ di numeri reali, con $\sigma \geq 0$, si denota con $\mathcal{N}(m, \sigma^2)$ l'immagine di $\mathcal{N}(0, 1)$ mediante l'applicazione lineare affine $x \mapsto \sigma x + m$ da \mathbb{R} in \mathbb{R} . Se X è una variabile aleatoria reale avente come legge $\mathcal{N}(0, 1)$, allora la variabile aleatoria $Y = \sigma X + m$ ha come legge $\mathcal{N}(m, \sigma^2)$, e si ha:

$$P[Y] = m, \quad \text{Var}[Y] = \sigma^2.$$

È quindi naturale chiamare $\mathcal{N}(m, \sigma^2)$ la *legge normale di media m e varianza σ^2* . La variabile aleatoria Y si chiama, invece, *variabile aleatoria gaussiana (di media m e varianza σ^2)*.

Per $\sigma = 0$, la legge normale si riduce alla legge della costante m (ossia alla misura di Dirac ε_m); per $\sigma > 0$, essa ammette come densità la funzione

$$y \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-m}{\sigma}\right)^2\right].$$

In ogni caso, la funzione caratteristica della legge $\mathcal{N}(m, \sigma^2)$ è la funzione:

$$\varphi(t) = \exp(itm - \sigma^2 t^2/2).$$

Come immediata conseguenza delle proprietà delle funzioni caratteristiche, inoltre, si ottiene che

$$\mathcal{N}(m_1, \sigma_1^2) \star \mathcal{N}(m_2, \sigma_2^2) = \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2).$$

Da questa formula si deduce che *la combinazione lineare di un numero finito di variabili aleatorie indipendenti e gaussiane è essa stessa una variabile aleatoria gaussiana*. Quest'osservazione suggerisce la seguente definizione:

(3.1.3) Definizione. Una famiglia $(X_i)_{i \in I}$ di variabili aleatorie reali su (Ω, \mathcal{A}, P) si dice una *famiglia gaussiana* se ogni X_i è gaussiana e se ogni combinazione lineare finita delle X_i è una variabile aleatoria gaussiana. Si dice poi che $(X_i)_{i \in I}$ è una famiglia gaussiana *centrata* se, per giunta, tutte le X_i sono centrate.

Per una famiglia gaussiana la nozione di indipendenza si riduce a quella di non correlazione. Più precisamente:

(3.1.4) Proposizione. Su (Ω, \mathcal{A}, P) sia $(X_i)_{i \in I}$ una famiglia gaussiana di variabili aleatorie. I fatti seguenti sono, allora, equivalenti:

- (a) $(X_i)_{i \in I}$ è una famiglia di variabili aleatorie indipendenti.
- (b) $(X_i)_{i \in I}$ è una famiglia di variabili aleatorie a due a due non correlate.

Dimostrazione. L'implicazione (a) \Rightarrow (b) è ovvia. Per quanto riguarda l'implicazione (b) \Rightarrow (a): senza ledere la generalità possiamo supporre che ogni X_i sia centrata. Inoltre, per il criterio fondamentale d'indipendenza, basta provare che, per ogni parte finita J di I , la famiglia $(X_i)_{i \in J}$ è indipendente. Per fare questo, supponiamo fissata una parte finita J di I : basta dimostrare che la funzione caratteristica del blocco $[X_i]_{i \in J}$ coincide con il prodotto tensore delle funzioni caratteristiche delle X_i . In effetti, fissato un elemento t in \mathbb{R}^J , poiché $(X_i)_{i \in J}$ è una famiglia gaussiana, la variabile aleatoria $\langle t, X \rangle$ ha legge normale e, denotata con σ_i^2 la varianza della variabile aleatoria X_i , poiché le variabili aleatorie X_i sono non correlate, essa ha come varianza il numero:

$$\tau^2 = \sum_{i \in J} t_i^2 \sigma_i^2.$$

La tesi segue, dunque, dall'uguaglianza seguente:

$$\varphi_X(t) = \varphi_{\langle t, X \rangle}(1) = \exp(-\tau^2/2) = \prod_{i \in J} \exp(-t_i^2 \sigma_i^2/2) = \prod_{i \in J} \varphi_{X_i}(t_i).$$

(3.1.5) Proposizione. Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie gaussiane sullo spazio probabilizzato (Ω, \mathcal{A}, P) . Per ciascun intero positivo n , sia $\mathcal{N}(m_n, \sigma_n^2)$ la legge di X_n . Se la successione X_n converge in \mathcal{L}^2 verso una variabile aleatoria X , allora X è gaussiana.

Dimostrazione. Si ponga:

$$m = \lim_n m_n, \quad \sigma^2 = \lim_n \sigma_n^2.$$

Per ipotesi è chiaro che la successione

$$\varphi_{X_n}(t) = P[e^{itX_n}] = \exp(itm_n - \sigma_n^2 t^2/2)$$

converge puntualmente verso la funzione

$$\varphi_X(t) = \exp(itm - \sigma^2 t^2/2).$$

Ciò basta per concludere che la successione (X_n) converge in legge verso una variabile aleatoria gaussiana di legge $\mathcal{N}(m, \sigma^2)$. Ora, poiché (X_n) converge verso X in \mathcal{L}^2 , la convergenza avviene anche in legge e dunque X ha necessariamente legge gaussiana $\mathcal{N}(m, \sigma^2)$.

3.2 Definizione di processo di Wiener

Sia assegnato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, una filtrazione \mathcal{F} avente \mathbb{R}_+ come insieme dei tempi.

(3.2.1) Definizione. Un processo X su (Ω, \mathcal{A}, P) , avente \mathbb{R}_+ come insieme dei tempi, è detto *additivo* (rispetto alla filtrazione \mathcal{F}) se è adattato a \mathcal{F} , e se, per ciascun numero reale positivo s , il blocco $[X_{s+t} - X_s]_{t \in \mathbb{R}_+}$ è indipendente dalla tribù \mathcal{F}_s .

Il seguente risultato è un utile criterio per verificare quando un processo adattato sia additivo.

(3.2.2) Proposizione. Su (Ω, \mathcal{A}, P) , sia X un processo adattato alla filtrazione \mathcal{F} , avente \mathbb{R}_+ come insieme dei tempi. Affinché X sia additivo è (necessario e) sufficiente che, per ogni coppia s, t di numeri reali positivi, sia indipendente da \mathcal{F}_s la variabile aleatoria $X_{t+s} - X_s$.

Dimostrazione. Si supponga fissato un numero reale positivo s . Si scelga una partizione finita

$$0 = t_0 < t_1 < \dots < t_n < \infty$$

e si consideri, per ogni fissato intero j compreso tra 1 e n , la variabile aleatoria

$$\Delta_j = X_{s+t_j} - X_{s+t_{j-1}}.$$

Comunque preso un elemento H in \mathcal{F}_s , è facile dimostrare, per induzione, che il blocco $[I_H, \Delta_1, \dots, \Delta_{j-1}]$ è indipendente dalla variabile aleatoria Δ_j . Per il criterio d'indipendenza per successioni è allora chiaro che $(I_H, \Delta_1, \dots, \Delta_n)$ è una famiglia finita di variabili aleatorie indipendenti. In particolare, la variabile aleatoria I_H risulta indipendente dal blocco $[\Delta_1, \dots, \Delta_n]$ e, per il principio della composizione, essa è indipendente dal blocco

$$[\Delta_1, \Delta_1 + \Delta_2, \dots, \Delta_1 + \dots + \Delta_n].$$

Ciò significa che la tribù \mathcal{F}_s è indipendente dal blocco $[X_{s+t_j} - X_s]_{1 \leq j \leq n}$ e quindi, per il criterio fondamentale d'indipendenza, essa risulta indipendente dal blocco $[X_{s+t} - X_s]_{t \in \mathbb{R}_+}$.

(3.2.3) Definizione. Un processo B su (Ω, \mathcal{A}, P) , avente \mathbb{R}_+ come insieme dei tempi e nullo in 0, è detto un *processo di Wiener* (o *moto browniano*) relativo alla filtrazione \mathcal{F} se è un processo additivo (rispetto a \mathcal{F}) con traiettorie continue e, per ogni coppia s, t di numeri reali positivi, $B_{s+t} - B_s$ ha legge $\mathcal{N}(0, t)$, ossia

$$(B_{s+t} - B_s)(P) = \mathcal{N}(0, t).$$

Osserviamo, innanzitutto, che l'ipotesi che B sia nullo in 0 è inessenziale: in effetti le proprietà alle quali deve sottostare un processo di Wiener riguardano soltanto gli incrementi e dunque non dipendono dalla variabile aleatoria iniziale. Inoltre, se B è un processo di Wiener, allora, per ogni coppia s, t di numeri reali positivi, la variabile aleatoria $B_{t+s} - B_s$ è indipendente da B_s . Più in generale, per ogni partizione finita $0 \leq t_1 < t_2 < \dots < t_n$,

$$(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$$

è una famiglia finita di variabili aleatorie gaussiane e indipendenti. In realtà questa proprietà vale anche quando si considerino dei tempi d'arresto. Più precisamente:

(3.2.4) Proposizione. Sia B un processo di Wiener rispetto alla filtrazione \mathcal{F} . Allora $(B_t)_{t \geq 0}$ è una martingala. Inoltre tale è il processo $(B_t^2 - t)_{t \in \mathbb{R}_+}$.

Dimostrazione. La prima affermazione è conseguenza delle proprietà elementari della speranza condizionale: si ha infatti, per ogni coppia s, t di numeri reali positivi,

$$P[B_{t+s} - B_s \mid \mathcal{F}_s] = P[B_{t+s} - B_s] = 0.$$

Per quanto riguarda l'altra affermazione, innanzitutto osserviamo che, per ogni coppia s, t di numeri reali positivi, si ha:

$$P[(B_{t+s} - B_s)^2 \mid \mathcal{F}_s] = P[B_{t+s}^2 + B_s^2 \mid \mathcal{F}_s] - 2B_s P[B_{t+s} \mid \mathcal{F}_s] = P[B_{t+s}^2 - B_s^2 \mid \mathcal{F}_s].$$

Ne segue, dunque, sfruttando il fatto che $B_{t+s} - B_s$ è indipendente dalla tribù \mathcal{F}_s , che ha legge $\mathcal{N}(0, t)$, e le proprietà dell'indipendenza condizionale,

$$P[(B_{t+s}^2 - (t+s)) - (B_s^2 - s) \mid \mathcal{F}_s] = P[(B_{t+s} - B_s)^2 \mid \mathcal{F}_s] - t = P[(B_{t+s} - B_s)^2] - t = 0$$

e anche la seconda affermazione è così provata.

(3.2.5) Proposizione. Supponiamo fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, un processo B nullo in 0 e con traiettorie continue. Le condizioni che seguono sono allora equivalenti:

(a) B è un processo di Wiener rispetto alla filtrazione naturale.

(b) B è una famiglia gaussiana centrata e, per ogni coppia s, t di numeri reali positivi, si ha:

$$(3.2.6) \quad P[B_s B_t] = s \wedge t.$$

Dimostrazione. Se B è un processo di Wiener, con un semplice ragionamento per induzione si prova che B è una famiglia gaussiana. Inoltre, per ogni coppia s, t di numeri reali positivi, si ha:

$$P[B_{t+s} B_s] = P[(B_{t+s} - B_s) B_s] + P[B_s^2] = P[B_{t+s} - B_s] P[B_s] + \text{Var}[B_s] = s.$$

Viceversa, si supponga che B sia una famiglia gaussiana e centrata di variabili aleatorie verificanti la relazione (3.2.6). Allora, per ogni coppia s, t di numeri reali positivi, la variabile aleatoria $B_{t+s} - B_s$ è gaussiana e centrata; inoltre, poiché

$$P[(B_{t+s} - B_s)^2] = P[B_{t+s}^2] + P[B_s^2] - 2P[B_{t+s} B_s] = t + s + s - 2s = t,$$

la variabile aleatoria $B_{t+s} - B_s$ ha legge $\mathcal{N}(0, t)$. Per concludere basta, dunque, verificare che, per ogni coppia s, t di numeri reali positivi, $B_{t+s} - B_s$ è indipendente dalla tribù $\mathcal{T}(B_u : u \leq s)$. Per vedere questo basta osservare (si veda (3.1.4)) che le variabili aleatorie $B_{t+s} - B_s$ e B_u sono non correlate, per ciascun numero reale u , con $u \leq s$. In effetti, vale la relazione:

$$\text{Cov}(B_{t+s} - B_s, B_u) = P[(B_{t+s} - B_s) B_u] = (t + s) \wedge u - s \wedge u = 0$$

e ciò basta per concludere.

(3.2.7) Esempio. Sia B un assegnato processo di Wiener sullo spazio probabilizzato (Ω, \mathcal{A}, P) (relativo alla propria filtrazione naturale \mathcal{F}). Fissato un numero reale positivo s , è facile riconoscere allora che il processo \tilde{B} , definito da

$$\tilde{B}_t = B_{t+s} - B_s$$

è un processo di Wiener relativo alla propria filtrazione naturale $\tilde{\mathcal{F}}$. Inoltre, non è difficile verificare che, per ogni numero reale positivo t , la tribù $\tilde{\mathcal{F}}_t$ è indipendente dalla tribù \mathcal{F}_s .

Non è difficile poi provare che, se T è un tempo d'arresto finito allora il processo \tilde{B} , definito da

$$\tilde{B}_t = B_{T+t} - B_T$$

è un processo di Wiener indipendente dalla tribù \mathcal{F}_S . A questo scopo, osserviamo innanzitutto che, senza ledere la generalità, possiamo supporre che T sia discreto. Nel caso generale basterà considerare una successione crescente (T_n) di tempi d'arresto discreti, ammettenti T come involucro superiore e ragionare su ciascun T_n . In queste ipotesi, osserviamo che, per ogni coppia α, β di numeri reali, denotata con D la parte numerabile di \mathbb{R}_+ costituita dai valori di T , si ha, tramite la formula di disintegrazione, e posto $H_s = \{T = s\}$,

$$\begin{aligned} P\{B_{T+t} - B_T > \alpha, B_T > \beta\} &= \sum_{s \in D} P_{H_s}\{B_{t+s} - B_s > \alpha, B_s > \beta\} P(H_s) \\ &= \sum_{s \in D} P_{H_s}\{B_{t+s} - B_s > \alpha\} P_{H_s}\{B_s > \beta\} P(H_s). \end{aligned}$$

Ora, osservato che $B_{t+s} - B_s$ è indipendente da H_s , e denotata con Φ la funzione di ripartizione della legge $\mathcal{N}(0, 1)$, si ha

$$\begin{aligned} P\{B_{T+t} - B_T > \alpha, B_T > \beta\} &= (1 - \Phi(\alpha/\sqrt{t})) \sum_{s \in D} P_{H_s}\{B_s > \beta\} P(H_s) \\ &= P\{B_{T+t} - B_T > \alpha\} P\{B_T > \beta\}, \end{aligned}$$

dove l'ultima eguaglianza segue dal fatto che la variabile aleatoria $B_{T+t} - B_T$ ha legge $\mathcal{N}(0, t)$ (come si riconosce facilmente utilizzando ancora la disintegrazione).

(3.2.8) Osservazione. Denotiamo con W lo spazio $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ (costituito da tutte le funzioni continue di \mathbb{R}_+ in \mathbb{R}), e con $[\xi_t]_{t \geq 0}$ il processo canonico su W , definito da $\xi_t(\omega) = \omega(t)$, per ogni numero reale positivo t , e ogni elemento ω di W . Inoltre, porremo

$$\mathcal{W} = \mathcal{T}(\xi_t : t \in \mathbb{R}_+).$$

Lo spazio misurabile (W, \mathcal{W}) è detto lo *spazio canonico di Wiener*. Un processo reale X , definito su un'opportuno spazio probabilizzato (Ω, \mathcal{A}, P) , avente traiettorie appartenenti a W , si può identificare con il blocco $[X_t]_{t \geq 0}$, *pensato come variabile aleatoria a valori in (W, \mathcal{W})* . Pertanto la sua legge sarà una misura di probabilità sulla tribù \mathcal{W} .

Se B è un processo di Wiener, su uno spazio probabilizzato (Ω, \mathcal{A}, P) , relativo alla propria filtrazione naturale, la sua legge si chiama *misura di Wiener* (sullo spazio (W, \mathcal{W})), la quale non dipende dal particolare processo di Wiener scelto. La si denoti con P_w .

Giova forse osservare che, se μ è una misura di probabilità su $\mathcal{B}(\mathbb{R})$, allora $\mu \star P_w$ è l'unica misura di probabilità secondo la quale il processo canonico sia un processo di Wiener avente μ come legge iniziale.

Da quest'osservazione segue che, per dimostrare l'effettiva esistenza della misura di Wiener, basta costruire, su di un opportuno spazio probabilizzato (Ω, \mathcal{A}, P) , un processo di Wiener rispetto alla propria filtrazione naturale.

3.3 Costruzione del processo di Wiener

Sia $(e_k)_{k \geq 1}$ una base ortonormale dello spazio di Hilbert $L^2(\lambda)$, dove λ denota la misura di Borel-Lebesgue sullo spazio misurabile $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Per ciascun indice k , si denoti con E_k un rappresentante di e_k . Si ponga:

$$G_k(t) = \int_0^t E_k(s) \, ds.$$

Fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) , sia $(X_k)_{k \geq 1}$ una successione di variabili aleatorie reali e indipendenti, isonome di legge $\mathcal{N}(0, 1)$. Per ciascun intero positivo n , si ponga

$$B^{(n)}(t, \omega) = \sum_{k=1}^n G_k(t) X_k(\omega).$$

Poiché (e_k) è una base ortonormale di $L^2(\lambda)$, la funzione $I_{[0,t]}$ è equivalente (modulo λ) alla serie $\sum_{k \geq 1} (\widetilde{I_{[0,t]}}, e_k) E_k$ (dove si denota con (\cdot, \cdot) il prodotto scalare dello spazio di Hilbert $L^2(\lambda)$, e con \widetilde{f} denota la classe di equivalenza, in $L^2(\lambda)$, avente f come rappresentante) e quindi si ha:

$$(3.3.1) \quad \sum_{k \geq 1} G_k(t)^2 = \sum_{k \geq 1} (\widetilde{I_{[0,t]}}, e_k)^2 = (\widetilde{I_{[0,t]}}, \widetilde{I_{[0,t]}}) = t.$$

Inoltre, poiché $(X_k)_{k \geq 1}$ è una successione di variabili aleatorie indipendenti ed isonome di legge $\mathcal{N}(0, 1)$, detto x_k un elemento di $L^2(P)$ avente come rappresentante la variabile aleatoria X_k , la successione (x_k) è una base ortonormale in $L^2(P)$. Ne segue che la successione

$$\sum_{k=1}^n G_k(t) x_k(\omega)$$

converge in norma L^2 (poiché converge la serie (3.3.1) dei coefficienti di Fourier) verso un elemento dello spazio $L^2(P)$ avente come rappresentante la variabile aleatoria

$$(3.3.2) \quad B(t, \omega) = \sum_{k \geq 1} G_k(t) X_k(\omega).$$

In particolare, la famiglia $(B_t)_{t \in \mathbb{R}_+}$ è una famiglia gaussiana e centrata di variabili aleatorie (si veda (3.1.5)). Inoltre, per ogni coppia s, t di numeri reali positivi, per l'uguaglianza di Parseval, si ha:

$$\begin{aligned} P[B_s B_t] &= \sum_{k \geq 1} G_k(t) G_k(s) \\ &= \sum_{k \geq 1} (\widetilde{I_{[0,t]}}, e_k) (\widetilde{I_{[0,s]}}, e_k) \\ &= \sum_{k \geq 1} (\widetilde{I_{[0,t]}}, \widetilde{I_{[0,s]}}) = t \wedge s. \end{aligned}$$

Per costruire un processo di Wiener, basta costruire, a partire da B , una modificazione \hat{B} di B avente traiettorie continue (si veda a questo proposito (3.2.5)). Questo segue come corollario di un celebre risultato dovuto a Kolmogorov.

(3.3.3) Teorema (di hölderianità di Kolmogorov). *Su un fissato spazio probabilizzato (Ω, \mathcal{A}, P) sia dato un processo reale X avente \mathbb{R}_+ come insieme dei tempi. Si supponga che esistano α, δ, C costanti reali strettamente positive, tali che, per ogni coppia s, t di numeri reali positivi, abbia luogo la relazione*

$$P[|X_s - X_t|^\alpha] \leq C|s - t|^{1+\delta}.$$

Allora esiste una modificazione di X con traiettorie localmente hölderiane di esponente γ , non appena γ sia minore di δ/α .

Allo scopo di provare questo teorema occorre premettere un lemma. Si denoti con D l'insieme costituito da tutti i numeri diadici, ossia da tutti i numeri della forma $k2^{-n}$, con k e n interi positivi. Fissato un intero positivo n , si denoti con D_n l'insieme dei numeri diadici di stadio n -esimo.

(3.3.4) Lemma. *Siano m un numero reale positivo, f una funzione reale definita su $D \cap [0, m]$ e γ un numero reale strettamente positivo. Si supponga che esista un intero positivo n_0 tale che, per ogni intero n superiore o eguale a n_0 , e ogni coppia s, t di elementi consecutivi di $D_n \cap [0, m]$, si abbia $|f(t) - f(s)| \leq 2^{-n\gamma}$.*

Allora f è hölderiana di esponente γ , ossia f è prolungabile ad una funzione hölderiana di esponente γ definita su $[0, m]$.

Dimostrazione. Mostriamo dapprima che esiste una costante reale C tale che, comunque si prendano un intero n superiore o eguale a n_0 , una coppia a, b di elementi consecutivi di $D_n \cap [0, m]$ e un elemento x di D tra essi compreso, risulti

$$|f(x) - f(a)| \leq C2^{-n\gamma}, \quad |f(x) - f(b)| \leq C2^{-n\gamma}.$$

A questo scopo, osserviamo che si ha:

$$x = a + \sum_{h \geq 1} \varepsilon_h 2^{-(n+h)},$$

dove $(\varepsilon_h)_{h \geq 1}$ è una successione di elementi di $\{0, 1\}$ definitivamente nulla. Dunque, ponendo, per ciascun intero strettamente positivo h ,

$$s_0 = a, \quad s_h = s_{h-1} + \varepsilon_h 2^{-(n+h)},$$

si definisce una successione crescente $(s_h)_{h \geq 0}$ di elementi di D i cui termini coincidono definitivamente con x . Inoltre, per ogni intero strettamente positivo h , i due termini s_{h-1} e s_h o coincidono o sono elementi consecutivi di D_{n+h} . Si ha, pertanto,

$$|f(x) - f(a)| \leq \sum_{h \geq 1} |f(s_h) - f(s_{h-1})| \leq \sum_{h \geq 1} 2^{-(n+h)\gamma} = C2^{-n\gamma},$$

con C costante reale. Un ragionamento analogo si applica a $|f(x) - f(b)|$.

Presi ora due elementi x, y di $D \cap [0, m]$, con

$$x < y, \quad |x - y| < 2^{-n_0},$$

si consideri l'intero n (superiore o eguale a n_0) caratterizzato dalla relazione

$$2^{-(n+1)} \leq |x - y| \leq 2^{-n}.$$

Esistono allora in $D_n \cap [0, m]$ due punti consecutivi a, b , con $a \leq x \leq y \leq b$, oppure tre punti consecutivi a, b, c , con $a \leq x \leq b \leq y \leq c$. In entrambi i casi, tuttavia, risulta

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f(b)| + |f(b) - f(y)| \\ &\leq 2C2^{-n\gamma} = K2^{-(n+1)\gamma} \\ &\leq K|x - y|^\gamma, \end{aligned}$$

con K costante. La funzione f è dunque uniformemente continua, ossia univocamente prolungabile in una funzione g continua su $[0, m]$. Inoltre, tale prolungamento è una funzione hölderiana di esponente γ perché verifica la relazione

$$|g(x) - g(y)| \leq K|x - y|^\gamma$$

per ogni coppia x, y di elementi di $[0, m]$ con $|x - y| \leq 2^{-n_0}$.

Veniamo ora alla dimostrazione del teorema. Per ogni intero positivo n , si denoti con Z_n l'involuppo superiore di tutte le variabili aleatorie della forma $|X_t - X_s|$, con s, t coppia di elementi consecutivi di $D_n \cap [0, m]$. Denotato con $A_{m,\gamma}$ l'evento costituito dagli elementi ω in Ω tali che la restrizione della traiettoria $X(\cdot, \omega)$ a $[0, m]$ sia hölderiana di esponente γ , l'insieme $\liminf_n \{Z_n \leq 2^{-n\gamma}\}$, grazie al lemma precedente, è contenuto in $A_{m,\gamma}$. Cominciamo col provare che $\liminf_n \{Z_n \leq 2^{-n\gamma}\}$ è quasi certo, ossia che l'evento $\limsup_n \{Z_n > 2^{-n\gamma}\}$ è trascurabile. Per il primo lemma di Borel–Cantelli, basta provare che la serie

$$(3.3.5) \quad \sum_{n \geq 0} \{Z_n > 2^{-n\gamma}\}$$

è convergente. In effetti, per ogni coppia s, t di elementi consecutivi di $D_n \cap [0, m]$, si ha (grazie alla disuguaglianza di Markov):

$$P\{|X_t - X_s| > 2^{-n\gamma}\} \leq 2^{n\gamma\alpha} P[|X_t - X_s|^\alpha] \leq 2^{n\gamma\alpha} C 2^{-n(1+\delta)}.$$

Ne segue che

$$P\{Z_n > 2^{-n\gamma}\} \leq m 2^n 2^{n\gamma\alpha} C 2^{-n(1+\delta)} = m C 2^{n(\gamma\alpha - \delta)}$$

e dunque, poiché risulta $\gamma < \delta/\alpha$, la serie (3.3.5) deve necessariamente convergere, per cui $\limsup_n \{Z_n > 2^{-n\gamma}\}$ è un evento trascurabile e, di conseguenza, $A_{m,\gamma}$ è quasi certo.

Si denoti con A l'intersezione di tutti gli eventi $A_{m,\gamma}$, con m intero positivo e γ numero razionale strettamente compreso tra 0 e δ/α . È chiaro che A è un evento quasi certo. Si ponga, pertanto,

$$\tilde{X}_t = I_A \liminf_{s \rightarrow t, s \in D} X_s.$$

Il processo $(\tilde{X}_t)_{t \in \mathbb{R}_+}$ ha traiettorie localmente hölderiane, per γ inferiore a δ/α , ed è una modificazione di X . Ciò conclude la dimostrazione.

(3.3.6) Corollario. *Il processo B definito in (3.3.2) ammette una modificazione \hat{B} con traiettorie hölderiane di esponente γ , per ciascun numero reale γ inferiore a $1/2$. Di conseguenza, \hat{B} è un processo di Wiener.*

Dimostrazione. Fissati un numero reale α strettamente positivo e una coppia s, t di numeri reali positivi, si ha:

$$P[|B_t - B_s|^\alpha] = C_\alpha |t - s|^{\alpha/2},$$

dove C_α denota il momento assoluto di ordine α della legge normale $\mathcal{N}(0, 1)$. Perciò, supposto $\alpha/2 > 1$, ossia $\alpha > 2$, se si applica il teorema (3.3.3) con $\delta = \frac{\alpha}{2} - 1$, si osserva che, per ogni numero reale γ verificante la relazione

$$0 < \gamma < \frac{1}{2} - \frac{1}{\alpha},$$

esiste una modificazione \hat{B} di B con traiettorie hölderiane di esponente γ . Osservando che i numeri γ , tali ch'esista α maggiore di 2 per i quali valga la relazione precedente, sono tutti i numeri γ strettamente compresi tra 0 e $1/2$, si deduce che è possibile trovare una modificazione \hat{B} di B con traiettorie localmente hölderiane di esponente γ strettamente compreso tra 0 e $1/2$. Per concludere basta osservare che \hat{B} verifica la condizione (b) di (3.2.5) e dunque è un processo di Wiener.

3.4 Proprietà delle traiettorie di un processo di Wiener

(3.4.1) Teorema. *Sia B un processo di Wiener definito sullo spazio probabilizzato (Ω, \mathcal{A}, P) . Allora l'insieme costituito dagli elementi ω di Ω per i quali la traiettoria $B(\cdot, \omega)$ ammette in qualche punto di \mathbb{R}_+ derivata destra inferiore e derivata destra superiore entrambe finite è contenuto in un evento trascurabile.*

Allo scopo di dimostrare questo teorema, cominciamo con la seguente semplice osservazione.

(3.4.2) Osservazione. Se, sullo spazio (Ω, \mathcal{A}, P) , è data una variabile aleatoria reale Z con legge normale $\mathcal{N}(0, 1)$ e se, per ogni numero reale positivo, si pone

$$F(x) = P\{|Z| \leq x\} = 2(2\pi)^{-1/2} \int_0^x \exp(-z^2/2) dz,$$

allora, al tendere di x a zero, si ha $F(x) \sim F'(0)x = 2(2\pi)^{-1/2}x$, e quindi $F(x) \approx x$.

Sfruttando questo risultato elementare, proviamo il lemma seguente.

(3.4.3) Lemma (di Erdős). Nelle ipotesi del teorema, poniamo, per ogni coppia n, k d'interi strettamente positivi,

$$J_k^n = [(k-1)/n, k/n[, \quad \Delta_k^n = |B_{k/n} - B_{(k-1)/n}|.$$

Inoltre, fissata una costante reale positiva c , poniamo

$$C_n = \bigcup_{k=1}^n \{ \Delta_{k+1}^n \vee \Delta_{k+2}^n \vee \Delta_{k+3}^n \leq c/n \}.$$

Allora l'evento $\liminf_n C_n$ è trascurabile.

Dimostrazione. Grazie al lemma di Fatou, basterà provare la relazione

$$(3.4.4) \quad \lim_n P(C_n) = 0.$$

A questo scopo, osserviamo che, per ogni coppia n, k d'interi strettamente positivi, $(\Delta_{k+1}^n, \Delta_{k+2}^n, \Delta_{k+3}^n)$ è una terna di variabili aleatorie indipendenti, tutte isonome a Δ_1^n . Si può dunque scrivere (tenendo conto dell'osservazione (3.4.2))

$$\begin{aligned} P(C_n) &\leq \sum_{k=1}^n P\{ \Delta_{k+1}^n \vee \Delta_{k+2}^n \vee \Delta_{k+3}^n \leq c/n \} \\ &= n(P\{ \Delta_1^n \leq c/n \})^3 \\ &= n(P\{ \sqrt{n}\Delta_1^n \leq c/\sqrt{n} \})^3 \approx n(1/\sqrt{n})^3 = 1/\sqrt{n}. \end{aligned}$$

La relazione (3.4.4) è così dimostrata.

Veniamo ora alla dimostrazione del teorema. Basterà provare che, per ogni costante reale positiva c , l'insieme

$$(3.4.5) \quad A = \bigcup_{s \in [0,1[} \{ \limsup_{t \rightarrow s, t > s} |B_t - B_s|/(t-s) < c \}$$

è contenuto in un evento trascurabile. Infatti, facendo variare c in \mathbb{N} se ne potrà dedurre che l'insieme costituito dagli ω di Ω per i quali la traiettoria $B(\cdot, \omega)$ ammette in qualche punto di $[0, 1[$ derivata destra inferiore e derivata destra superiore entrambe finite è contenuto in un evento trascurabile. Applicando poi questo risultato, per ogni intero positivo m , al processo di Wiener $(B_{m+t} - B_m)_{t \geq 0}$, si potrà ottenere la tesi del teorema.

Poniamo, per ogni elemento s di $[0, 1[$, e ogni intero n strettamente positivo,

$$\begin{aligned} A_s &= \{ \limsup_{t \rightarrow s, t > s} |B_t - B_s|/(t-s) < c \}, \\ D_{n,s} &= \bigcap_{t \in [s, s+(4/n)]} \{ |B_t - B_s| < c(t-s) \}. \end{aligned}$$

Osserviamo che, per ogni elemento ω di $D_{n,s}$, la traiettoria $B(\cdot, \omega)$ ha, sull'intervallo $[s, s + (4/n)]$, un'oscillazione non superiore a $(8c)/n$. Poniamo inoltre

$$(3.4.6) \quad D_n = \bigcup_{s \in [0,1[} D_{n,s}.$$

Si ha allora, per ogni s ,

$$A_s \subset \liminf_n D_{n,s} \subset \liminf_n D_n.$$

Dalla relazione (3.4.5) discende dunque

$$(3.4.7) \quad A = \bigcup_{s \in [0,1[} A_s \subset \liminf_n D_n.$$

D'altra parte, grazie alla relazione (3.4.6) di D_n , si ha, con le notazioni del lemma,

$$D_n = \bigcup_{k=1}^n \bigcup_{s \in J_k^n} D_{n,s} \subset \bigcup_{k=1}^n P\{\Delta_{k+1}^n \vee \Delta_{k+2}^n \vee \Delta_{k+3}^n \leq (8c)/n\},$$

dove l'inclusione finale è dovuta al fatto che, se ω è un elemento di $D_{n,s}$, con $s \in J_k^n$, allora ciascuno dei tre intervalli $J_{k+1}^n, J_{k+2}^n, J_{k+3}^n$ risulta contenuto nell'intervallo $[s, s + (4/n)]$, sul quale la traiettoria $B(\cdot, \omega)$ ha un'oscillazione non superiore a $(8c)/n$. Dalla relazione (3.4.7) discende dunque che l'insieme A è contenuto nell'evento

$$\liminf_n \bigcup_{k=1}^n P\{\Delta_{k+1}^n \vee \Delta_{k+2}^n \vee \Delta_{k+3}^n \leq (8c)/n\}.$$

Poiché questo, grazie al lemma precedente, è un evento trascurabile, il teorema è dimostrato.

(3.4.8) Teorema. *Sia B un processo di Wiener definito sullo spazio probabilizzato (Ω, \mathcal{A}, P) . Allora:*

(a) *Per quasi ogni ω , la traiettoria $B(\cdot, \omega)$ non è monotona su alcun intervallo proprio (e quindi possiede un insieme denso di punti di massimo locale).*

(b) *Per quasi ogni ω , la traiettoria $B(\cdot, \omega)$ assume massimi diversi su ogni coppia d'intervalli propri, compatti e disgiunti, con estremi razionali (e quindi non possiede alcun punto di massimo locale che sia stretto).*

Dimostrazione. L'affermazione (a) è una conseguenza del teorema precedente. In effetti, se una traiettoria di B fosse monotona su un intervallo proprio, allora essa sarebbe dotata di derivata finita in quasi ogni punto di questo intervallo. Per provare l'affermazione (b) basta provare che, dati due intervalli compatti $H = [a, b]$ e $K = [c, d]$, con $0 \leq a < b < c < d < \infty$, e posto

$$M_H = \sup_{t \in H} B_t, \quad M_K = \sup_{t \in K} B_t,$$

l'evento $\{M_H = M_K\}$ è trascurabile. Per questo, basta scrivere l'evento in questione nella forma $\{M_H - B_b = M_K - B_b\}$ e osservare che la variabile aleatoria $M_K - B_b$ è indipendente dal blocco $[B_t]_{0 \leq t \leq b}$ (dunque da $M_H - B_b$) e dotata di densità, in quanto può esser messa nella forma

$$M_K - B_b = (B_c - B_b) + \sup_{c \leq t \leq d} (B_t - B_c)$$

e il primo termine è dotato di densità (legge della variabile aleatoria in questione sarà dunque il prodotto di convoluzione delle leggi delle variabili aleatorie di cui è somma). Ora, poiché l'evento $\{M_H = M_K\}$ corrisponde a dire che il blocco $[M_H - B_b, M_K - B_b]$ sta sulla diagonale, la tesi segue dal fatto che la diagonale è trascurabile secondo la misura di Lebesgue.

3.5 Il processo di Wiener multidimensionale

Supponiamo assegnato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, una filtrazione $(\mathcal{F}_t)_{t \geq 0}$.

(3.5.1) Definizione. Siano B^1, \dots, B^d processi di Wiener relativi alla filtrazione $(\mathcal{F}_t)_{t \geq 0}$. Si supponga che i blocchi $[B_t^1]_{t \geq 0}, \dots, [B_t^d]_{t \geq 0}$ siano variabili aleatorie indipendenti. Si dice allora che il processo B (sullo spazio (Ω, \mathcal{A}, P) , a valori in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$), definito da,

$$B_t = [B_t^1, \dots, B_t^d],$$

è un *processo di Wiener d -dimensionale* (relativo alla filtrazione $(\mathcal{F}_t)_{t \geq 0}$).

(3.5.2) Osservazione. Grazie all'indipendenza delle “componenti” del processo B , è immediato riconoscere che il processo B verifica le proprietà seguenti:

(a) Il processo B è adattato a $(\mathcal{F}_t)_{t \geq 0}$, e per ogni coppia s, t di numeri reali positivi, la variabile aleatoria $B_{s+t} - B_s$ è indipendente dalla tribù \mathcal{F}_s .

(b) Per ogni coppia s, t di numeri reali positivi, la legge su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, della variabile aleatoria $B_{s+t} - B_s$ coincide con la misura prodotto $\mathcal{N}(0, t)^{\otimes d}$. In particolare questa legge è dotata di densità (rispetto alla misura di Borel-Lebesgue) n_t , definita da,

$$n_t(x) = (2\pi t)^{-n/2} \exp(-|x|^2/(2t)).$$

(c) Il processo B ha traiettorie continue.

È facile riconoscere che queste tre proprietà caratterizzano completamente un processo di Wiener di dimensione d , nel senso che, se B è un processo a valori nello spazio misurabile $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, e verifica le proprietà (a), (b), (c) dell'osservazione precedente, allora esso è un processo di Wiener d -dimensionale.

IV

L'integrale stocastico

4.1 Convergenze negli spazi dei processi

Fissiamo uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, una filtrazione \mathcal{F} avente come insieme dei tempi \mathbb{R}_+ e verificante le condizioni abituali.

Denotiamo con \mathcal{R} lo spazio vettoriale costituito da tutti i processi regolari, continui a destra e adattati. Si denoti inoltre con \mathcal{R}_0 il sottospazio vettoriale di \mathcal{R} costituito da tutti i processi in \mathcal{R} nulli in 0. Denotiamo inoltre con $\tilde{\mathcal{R}}$ (resp. $\tilde{\mathcal{R}}_0$) lo spazio quoziente di \mathcal{R} (resp. \mathcal{R}_0) rispetto alla relazione d'indistinguibilità (cioè lo spazio quoziente di \mathcal{R} rispetto al sottospazio vettoriale costituito da tutti gli elementi evanescenti di \mathcal{R}).

(4.1.1) Osservazione. È facile riconoscere che ogni processo X in \mathcal{R} si può decomporre come somma di un processo nullo in 0 (ossia appartenente a \mathcal{R}_0) e di un processo “costante”. Basta osservare, a questo scopo, che si ha

$$(4.1.2) \quad X = \bar{X} + X_0,$$

dove

$$\bar{X}(t, \omega) = X(t, \omega) - X(0, \omega), \quad X_0(t, \omega) = X(0, \omega).$$

Se X appartiene a \mathcal{R} , denotiamo con X^* il processo definito come:

$$X_t^* = \sup_{0 \leq s \leq t} |X_s| = \sup_{r \in \mathbb{Q}_+, r \leq 1} |X_{rt}|.$$

Si vede facilmente che X^* è un processo regolare, continuo a destra e adattato, cioè X^* appartiene a \mathcal{R} .

(4.1.3) Definizione. Sia X un elemento di \mathcal{R} e sia (X^n) una successione di elementi di \mathcal{R} . Si dice allora che (X^n) converge verso X in \mathcal{R} se, per ogni numero reale positivo t , la successione di variabili aleatorie $(X^n - X)_t^*$ converge verso 0 in probabilità.

Nel seguito considereremo sempre lo spazio \mathcal{R} come munito della nozione di convergenza appena introdotta; tale nozione può essere chiamata *convergenza compatta in probabilità*. Essa induce, tramite passaggio al quoziente, una nozione di convergenza sullo spazio $\tilde{\mathcal{R}}$.

È facile riconoscere che, affinché una successione (X^n) di elementi di \mathcal{R} converga verso un elemento X di \mathcal{R} nel senso della definizione (4.1.3), è sufficiente che la convergenza in probabilità di $(X^n - X)_t^*$ abbia luogo per ciascun t appartenente a \mathbb{N} . Ne discende il seguente criterio di convergenza.

(4.1.4) Proposizione. Sia X un elemento di \mathcal{R} e sia (X^n) una successione di elementi di \mathcal{R} . Le condizioni che seguono sono, allora, equivalenti:

(a) La successione (X^n) converge verso X in \mathcal{R} .

(b) Da tutte le successioni estratte da $(X^n - X)^*$ è possibile estrarre una sottosuccessione che converge puntualmente verso 0 fuori da un insieme evanescente.

Da questa proposizione risultano i due corollari seguenti:

(4.1.5) Corollario. Sia (X^n) una successione di elementi di \mathcal{R} e sia (T_k) una successione di tempi d'arresto, con $\sup_k T_k = \infty$. Si supponga che, per ciascun k fissato, la successione $(X^{n|T_k})_n$ converga verso 0 in \mathcal{R} . Anche la successione (X^n) converge allora verso 0 in \mathcal{R} .

Dimostrazione. A meno di passare ad una sottosuccessione, si può supporre che esista un insieme evanescente G tale che, per ogni k fissato, la successione $((X^{n|T_k})^*)_n$ converga puntualmente verso 0 sul complementare di G ; la stessa cosa vale allora per la successione $((X^n)^*)$.

(4.1.6) Corollario. Sia (X^n) una successione di elementi di \mathcal{R} convergente verso 0 in \mathcal{R} . Allora, per ogni variabile aleatoria reale e positiva T , la successione (X_T^n) converge verso 0 in probabilità.

Dimostrazione. Fissati, a questo proposito, una coppia di ε, δ di numeri reali strettamente positivi, ed una variabile aleatoria reale e positiva T , esiste, per ipotesi, un intero positivo n_ε , tale che, per ogni intero n superiore a n_ε , si abbia

$$P\{|X_T^n| > \delta\} = \lim_k P\{|X_T^n| > \delta, T \leq k\} \leq P\{(X^n)_k^* > \delta\} \leq \varepsilon.$$

La proposizione che segue mostra che la convergenza introdotta in (4.1.3) proviene da una topologia di spazio metrico completo.

(4.1.7) Proposizione. *Lo spazio \mathcal{R} , munito della nozione di convergenza introdotta in (4.1.3), è uno spazio metrico completo.*

Dimostrazione. Dati due processi X, Y in \mathcal{R} , si ponga

$$d(X, Y) = \sum_{k \geq 0} 2^{-k} P[1 \wedge (X - Y)_k^*].$$

È chiaro che d è una distanza su \mathcal{R} . La nozione di convergenza introdotta da d equivale a quella definita in (4.1.3). Infatti, se (X^n) converge verso 0 in \mathcal{R} , allora, a meno di passare ad una sottosuccessione, possiamo supporre che $(X^n)^*$ converga verso 0 puntualmente al di fuori di un insieme evanescente G . Inoltre, a meno di sostituire X^n con $I_G X^n$, possiamo supporre $G = \emptyset$. Si ha allora

$$\lim_n d(X^n, 0) = \lim_n \sum_{k \geq 0} 2^{-k} P[1 \wedge (X^n)_k^*] = 0$$

per il teorema di convergenza dominata. Viceversa, si supponga che (X^n) converga verso 0 nella topologia indotta da d : si ha allora

$$\lim_n \sum_{k \geq 0} 2^{-k} P[1 \wedge (X^n)_k^*] = 0$$

e, dunque, la successione numerica $(P[1 \wedge (X^n)_k^*])$ converge verso 0. Ne segue che, per ogni numero reale strettamente positivo δ , per la disuguaglianza di Markov, si ha

$$P\{(X^n)_k^* > \delta\} \leq \delta^{-1} P[1 \wedge (X^n)_k^*]$$

ed il secondo membro converge verso 0 al tendere di k all'infinito. Dunque $((X^n)_k^*)$ converge verso 0 in probabilità o, ciò ch'è lo stesso, (X^n) converge verso 0 in \mathcal{R} .

Per concludere basta dimostrare la distanza d rende lo spazio \mathcal{R} completo. A questo scopo, sia (X^n) una successione di Cauchy di elementi di \mathcal{R} . Si può estrarre allora una sottosuccessione (X^{n_s}) tale che, per ciascun intero positivo s si abbia

$$d(X^{n_s}, X^{n_{s+1}}) = \sum_{k \geq 0} 2^{-k} P[1 \wedge (X^{n_s} - X^{n_{s+1}})_k^*] < 2^{-s}.$$

Si ponga inoltre

$$X = X^{n_0} + \sum_{s \geq 0} (X^{n_s} - X^{n_{s+1}}).$$

È chiaro che la successione (X^{n_s}) converge in \mathcal{R} verso X . La conclusione segue dunque dal fatto che (X^n) è una successione di Cauchy avente una sottosuccessione convergente.

(4.1.8) Teorema (principio di ricolloamento). Sia (X^k) una successione di elementi di \mathcal{R} , e sia (T_k) una successione di tempi d'arresto, con $\sup_k T_k = \infty$. Si supponga che, per ciascun k , il processo X^k sia indistinguibile da $X^{k+1}|_{T_k}$. Esiste allora un elemento Y (unico a meno di elementi indistinguibili) di \mathcal{R} tale che, per ciascun k , $Y|_{T_k}$ sia indistinguibile da X^k . In particolare, tale Y è prevedibile se tutti i processi X^k lo sono.

Dimostrazione. Per ipotesi esiste un elemento A di \mathcal{A} , trascurabile secondo P , tale che, per ciascun k , i due processi X^k e $X^{k+1}|_{T_k}$ coincidono sul complementare dell'insieme (prevedibile ed evanescente):

$$G =]0, \infty[\times A.$$

A meno di sostituire X^k con $X^k I_{G^c}$ possiamo supporre che l'insieme A sia vuoto. Dunque, posto $T_0 = 0$, il processo

$$Y = \sum_{k \geq 1} X^k I_{]T_{k-1}, T_k]}$$

verifica la desiderata proprietà.

Introduciamo ora un secondo spazio di processi. Per fare questo occorre premettere la definizione seguente:

(4.1.9) Definizione. Un processo X si dice *localmente limitato* se esiste una successione crescente (T_k) di tempi d'arresto, con $\sup_k T_k = \infty$, tale che ciascuno dei processi $X I_{[0, T_k]}$ sia limitato.

Denotiamo con \mathcal{P}_b (risp. con \mathcal{P}) lo spazio vettoriale costituito da tutti i processi prevedibili e limitati (risp. localmente limitati). Denotiamo poi con \mathcal{P}_0 il sottospazio vettoriale di \mathcal{P} costituito da tutti i processi in \mathcal{P} nulli in 0.

(4.1.10) Osservazione. Analogamente a quanto visto per i processi in \mathcal{R} , ogni processo H appartenente a \mathcal{P} si può decomporre in un processo nullo in 0 (ossia appartenente a \mathcal{P}_0) ed un processo costante. Con le notazioni introdotte in (4.1.1), tale decomposizione si può scrivere come

$$H = \bar{H} + H_0, \quad \text{con } \bar{H} \in \mathcal{P}_0.$$

È facile riconoscere che lo spazio \mathcal{P} contiene tutti i processi della forma $I_{[0, T]}$, con T tempo d'arresto. Inoltre esso contiene tutti i processi prevedibili elementari (non necessariamente nulli in 0), ossia quei processi della forma

$$V_0 + \sum_{i=1}^k I_{[s_i, t_i]} \otimes V_i,$$

dove s_i, t_i sono coppie numeri reali positivi, con $s_i < t_i$, V_i sono variabili aleatorie reali limitate e misurabili rispetto a \mathcal{F}_{s_i} e V_0 è misurabile rispetto a \mathcal{F}_0 .

(4.1.11) Proposizione. *Sia H un processo continuo a sinistra e adattato. Si supponga che H abbia traiettorie localmente limitate (ossia limitate su ogni intervallo limitato). Allora H appartiene a \mathcal{P} .*

Dimostrazione. Senza ledere la generalità possiamo supporre che H sia nullo in 0. Basta provare allora che H è un processo prevedibile. Si denoti con H^* il processo (avente $]0, \infty[$ come insieme dei tempi) definito da

$$H_t^* = \sup_{0 < s \leq t} |H_s|.$$

Poiché il processo H è prevedibile, tale è il processo H^* . Si ponga, per ciascun intero positivo k ,

$$T_k(\omega) = \inf\{t : H_t^*(\omega) > k\}.$$

È facile riconoscere che si tratta di un tempo d'arresto; inoltre, T_k è caratterizzato dalla relazione

$$\llbracket T_k, \infty \rrbracket = \{H^* > k\}$$

per ciascun intero positivo k . Tale relazione mostra che la successione (T_k) è crescente e che si ha $\sup_k T_k = \infty$. Inoltre, il processo

$$HI_{\llbracket 0, T_k \rrbracket} = HI_{\{H^* < k\}}$$

è limitato. Ne segue che il processo H è localmente limitato e quindi appartiene a \mathcal{P}_0 .

(4.1.12) Corollario. *Sia X un processo appartenente a \mathcal{R} . Allora il processo X_- appartiene a \mathcal{P} .*

(4.1.13) Definizione. Sia H un elemento di \mathcal{P} e sia (H^n) una successione di elementi di \mathcal{P} . Si dice che (H^n) converge verso H in \mathcal{P} se valgono le due condizioni seguenti:

- (a) (H^n) converge verso H puntualmente al di fuori di un insieme evanescente.
- (b) Esiste un elemento K di \mathcal{P} verificante la relazione $\sup_n |H^n| \leq K$ al di fuori di un insieme evanescente.

Nel seguito considereremo sempre lo spazio \mathcal{P} come munito della nozione di convergenza appena introdotta.

(4.1.14) Proposizione. *Sia X un elemento di \mathcal{R} e sia (X^n) una successione di elementi di \mathcal{R} convergente verso X in \mathcal{R} . La successione (X_-^n) (di elementi di \mathcal{P}) ammette allora una sottosuccessione convergente verso X_- in \mathcal{P} .*

Dimostrazione. Senza ledere la generalità possiamo supporre $X = 0$. Inoltre, a meno di passare ad una sottosuccessione, possiamo supporre che esista un elemento A di \mathcal{A} , trascurabile secondo P , tale che la successione $((X^n)^*)$ converga puntualmente verso 0 sul complementare dell'insieme (prevedibile ed evanescente):

$$G =]0, \infty[\times A.$$

A meno di sostituire X^n con $X^n I_{G^c}$ possiamo supporre che l'insieme A sia vuoto. La successione (X_-^n) converge allora verso 0 in ogni punto di $]0, \infty[$. Inoltre, il processo

$$Y = \sup_n (X^n)^*$$

è un processo appartenente a \mathcal{R} e tale che, per ogni intero positivo n , si abbia $|X_-^n| \leq Y_-$. Osservato che il processo Y_- appartiene allo spazio \mathcal{P} , questo basta per concludere.

4.2 La nozione di semimartingala e d'integrale stocastico

(4.2.1) Definizione. Un elemento X appartenente a \mathcal{R}_0 si chiama una *semimartingala* (nulla in 0) se esiste un'applicazione lineare f da \mathcal{P} a $\tilde{\mathcal{R}}_0$ che verifica le seguenti proprietà:

- (a) per ogni tempo d'arresto T , il processo $X^{|T}$ è un rappresentante di $f(I_{[0, T]})$;
- (b) (*proprietà di continuità*) per ogni successione (H^n) di elementi di \mathcal{P} convergente verso 0 in \mathcal{P} , la successione $(f(H^n))$ converge verso 0 in $\tilde{\mathcal{R}}_0$ o, ciò ch'è lo stesso, in $\tilde{\mathcal{R}}$.

(4.2.2) Lemma. Sia f un'applicazione lineare da \mathcal{P} a $\tilde{\mathcal{R}}_0$ che verifica la proprietà di continuità (4.2.1) (b) e che si annulla sulle indicatori degli insiemi della forma $[0, T]$, con T tempo d'arresto. L'applicazione f è allora identicamente nulla.

Dimostrazione. Il nucleo di f è uno spazio vettoriale monotono. Poiché esso contiene la classe delle indicatori degli insiemi della forma $[0, T]$, con T tempo d'arresto, esso contiene (per il teorema delle classi monotone) anche lo spazio \mathcal{P}_b di tutti i processi prevedibili e limitati. È sufficiente, dunque, ricordare che, per definizione, tutti gli elementi positivi H di \mathcal{P} sono involuppo superiore di una successione crescente di elementi positivi di \mathcal{P}_b (della forma $HI_{[0, T_k]}$). Ciò basta per concludere che il nucleo di f coincide con \mathcal{P} .

(4.2.3) Definizione. Sia X una semimartingala nulla in 0. Esiste allora, per il lemma precedente, un'unica applicazione lineare f da \mathcal{P} a $\tilde{\mathcal{R}}_0$ che verifica le proprietà (a) e (b) della definizione (4.2.1): tale applicazione è chiamata *integrale stocastico* associato alla semimartingala X . Il suo valore nel punto H è denotato con $H \bullet X$ (oppure con $\int H dX$ o anche $\int H_t dX_t$) e si chiama *integrale stocastico di H rispetto alla semimartingala X* .

(4.2.4) Definizione. Un elemento $X = \bar{X} + X_0$ appartenente a \mathcal{R} si chiama una *semimartingala*, se il processo \bar{X} (appartenente a \mathcal{R}_0) è una semimartingala nulla in 0. In questo caso, si chiama *integrale stocastico* associato alla semimartingala X l'applicazione lineare (da \mathcal{P} a $\tilde{\mathcal{R}}$) definita da

$$(4.2.5) \quad H \bullet X = H_0 X_0 + \bar{H} \bullet \bar{X}.$$

In questo caso, poi, si denota con $\int H dX$ la classe $\bar{H} \bullet \bar{X}$.

Applicando il lemma (4.2.2), è immediato provare il risultato seguente:

(4.2.6) Proposizione. Sia X una semimartingala. Allora l'applicazione $f : H \mapsto H \bullet X$, definita da (4.2.5), è l'unica applicazione lineare, da \mathcal{P} a $\tilde{\mathcal{R}}$ che verifica le due proprietà seguenti:

- (a) Per ogni tempo d'arresto T , il processo $X|_T$ è un rappresentante di $f(I_{[0,T]})$.
- (b) (proprietà di continuità) Per ogni successione (H^n) di elementi di \mathcal{P} convergente verso 0 in \mathcal{P} , la successione $(f(H^n))$ converge verso 0 in $\tilde{\mathcal{R}}$.

(4.2.7) Osservazioni. Innanzitutto osserviamo che l'insieme costituito da tutte le semimartingale è un sottospazio vettoriale di \mathcal{R} e che l'applicazione $(H, X) \mapsto H \bullet X$ è bilineare.

Inoltre, se X è un elemento evanescente di \mathcal{R} , allora X è una semimartingala e l'integrale stocastico ad essa associato è l'applicazione identicamente nulla. Per questo, due semimartingale indistinguibili definiscono lo stesso integrale stocastico. Di conseguenza, se Z è una classe di semimartingale a due a due indistinguibili, e H è un elemento di \mathcal{P} , si potrà denotare con $H \bullet Z$ (senza rischio di confusione) l'integrale stocastico di H rispetto ad un qualunque rappresentante della classe Z .

Infine, sia X una semimartingala. Allora, per ogni elemento H evanescente di \mathcal{P} , si ha $H \bullet X = 0$. Per provarlo basta, infatti, applicare la proprietà di continuità alla successione costante $H^n = 0$.

Sia X una semimartingala e sia H un elemento di \mathcal{P} . Siano S, T una coppia di tempi d'arresto finiti, con $S \leq T$. Si denota, allora, con

$$\int_S^T H dX \quad \text{oppure} \quad \int_S^T H_t dX_t$$

la classe d'equivalenza determinata dalle variabili aleatorie della forma $Y_T - Y_S$ dove Y è un rappresentante di $H \bullet X$.

(4.2.8) Proposizione (proprietà associativa). Sia X una semimartingala e sia H un elemento di \mathcal{P} . Allora, ogni rappresentante della classe $H \bullet X$ è una semimartingala e si ha

$$K \bullet (H \bullet X) = (KH) \bullet X$$

per ogni elemento K di \mathcal{P} .

Allo scopo di provare la proposizione dimostriamo prima un caso particolare:

(4.2.9) Lemma. *Sia X una semimartingala e sia T un tempo d'arresto. Allora, per ogni elemento H di \mathcal{P} ,*

$$(4.2.10) \quad (H \bullet X)^{|T} = (I_{\llbracket 0, T \rrbracket} H) \bullet X,$$

dove il primo membro denota l'elemento di $\tilde{\mathcal{R}}$ determinato dai processi della forma $Y^{|T}$, con Y rappresentante di $H \bullet X$.

Dimostrazione. Innanzitutto osserviamo che, senza ledere la generalità, possiamo supporre che X sia una semimartingala nulla in 0. Consideriamo l'applicazione g che, ad ogni elemento H di \mathcal{P} associa la differenza tra il primo membro e il secondo membro di (4.2.10). Si riconosce facilmente che, per ogni tempo d'arresto S , in virtù di (4.2.1)(a), si ha

$$g(I_{\llbracket 0, S \rrbracket}) = (X^{|S})^{|T} - X^{|S \wedge T} = 0.$$

È chiaro, inoltre, che g verifica la condizione (b) di (4.2.1). Ne segue dunque, dal lemma (4.2.2), che g è identicamente nulla e ciò basta per concludere.

Veniamo ora alla dimostrazione della proposizione: anche in questo caso ci si può ricondurre a considerare una semimartingala X nulla in 0. Consideriamo l'applicazione lineare f da \mathcal{P} a $\tilde{\mathcal{R}}$ definita da

$$f(K) = (KH) \bullet X.$$

Essa possiede chiaramente la proprietà di continuità (4.2.1) (b). Inoltre, se si denota con Y un rappresentante di $H \bullet X$, il lemma precedente mostra che, per ogni tempo d'arresto T , il processo $Y^{|T}$ è un rappresentante di

$$(I_{\llbracket 0, T \rrbracket} H) \bullet X = f(I_{\llbracket 0, T \rrbracket}).$$

Per le definizioni (4.2.1) e (4.2.3) ciò significa che Y è una semimartingala e che f è l'integrale stocastico associato a Y . La proposizione è dunque dimostrata.

La proposizione seguente esprime il carattere locale della nozione di semimartingala.

(4.2.11) Proposizione. *Sia X un processo appartenente a \mathcal{R} . Si supponga che esista una successione crescente (T_k) di tempi d'arresto, con $\sup_k T_k = \infty$, tali che, per ogni k , il processo arrestato $X^{|T_k}$ sia una semimartingala. Il processo X è allora una semimartingala.*

Dimostrazione. Sia dato un elemento H di \mathcal{P} . Si scelga, per ciascun k , una versione Y^k dell'integrale stocastico $H \bullet (X^{|T_k})$. Utilizzando la relazione

$$X^{|T_k} = (X^{|T_{k+1}})^{|T_k}$$

e la proprietà associativa (4.2.8), si vede facilmente che, per ogni k , il processo Y^k è indistinguibile da $Y^{k+1}|^{T_k}$: in effetti, ragionando “per classi monotone”, basta provarlo per H della forma $I_{[0,S]}$, con S tempo d’arresto. In questo caso si ha

$$(I_{[0,S]} \bullet (X^{T_{k+1}}))^{T_k} = (X^{T_{k+1} \wedge S})^{T_k} = X^{T_k \wedge S}.$$

Il principio di ricolloamento (4.1.8) mostra allora ch’esiste un elemento Z di \mathcal{R} tale che, per ogni k , il processo Z^{T_k} sia indistinguibile da Y^k (e dunque, sia un rappresentante di $H \bullet (X^{T_k})$). La classe di equivalenza di $\tilde{\mathcal{R}}$ della quale Z è un rappresentante non dipende che da H : la denoteremo, dunque, con $f(H)$. Abbiamo quindi definito un’applicazione lineare f da \mathcal{P} a $\tilde{\mathcal{R}}$. Questa applicazione possiede la proprietà (4.2.6)(a) poiché, per ogni tempo d’arresto T , e ogni intero positivo k , si ha

$$f(I_{[0,T]})^{T_k} = I_{[0,T]} \bullet (X^{T_k}) = X^{T_k \wedge T},$$

da cui, passando al limite per k tendente all’infinito, si trae $f(I_{[0,T]}) = X^{T^+}$. Proviamo che f possiede anche la proprietà (b). Sia (H^n) una successione di elementi di \mathcal{P} convergente verso 0 in \mathcal{P} . Per ciascun n sia Z^n un rappresentante di $f(H^n)$. Fissato k , il processo $Z^n|^{T_k}$ è allora un rappresentante di $H^n \bullet (X^{T_k})$. Di conseguenza, la successione $(Z^n|^{T_k})$ converge verso 0 in \mathcal{R} . Poiché tale conclusione è vera per ogni k , ne segue da (4.1.5) che la successione (Z^n) converge verso 0 in \mathcal{R} e questo basta per concludere che f è l’integrale stocastico associato alla semimartingala X .

La proposizione seguente mostra che, se si sostituisce \mathcal{P} con \mathcal{P}_b nella definizione di semimartingala, si ottiene una definizione equivalente.

(4.2.12) Proposizione. *Sia X un processo appartenente a \mathcal{R} . Si supponga che esista un’applicazione lineare f da \mathcal{P}_b a $\tilde{\mathcal{R}}$, verificante le proprietà seguenti:*

- (a) *Per ogni tempo d’arresto limitato T , X^{T^+} è un rappresentante di $f(I_{[0,T]})$.*
- (b) *Per ogni successione (H^n) di elementi di \mathcal{P}_b , convergente puntualmente verso 0 fuori da un insieme evanescente, e tale ch’esista una costante reale positiva c verificante la relazione $\sup_n |H^n| \leq c$ fuori da un insieme evanescente, la successione $f(H^n)$ converge verso 0 in $\tilde{\mathcal{R}}$.*

Il processo X è allora una semimartingala, e si ha $f(H) = H \bullet X$ per ogni elemento H di \mathcal{P}_b .

Dimostrazione. Innanzitutto osserviamo che l’ipotesi (b) implica che f si annulla su tutti gli elementi evanescenti di \mathcal{P}_b . La stessa ipotesi mostra inoltre che la proprietà (a) ha luogo anche per un tempo d’arresto T non limitato: basta applicare (b) alla successione (H^n) definita da $H^n = I_{[0,T]} - I_{[0,T \wedge n]}$. Ricordiamo che, dato un tempo d’arresto S qualunque, ogni elemento H di \mathcal{P}_b possiede la proprietà seguente: Se Y è un rappresentante di $f(H)$, Y^{S^+} è un rappresentante di $f(I_{[0,S]}H)$. Infatti, se si denota con \mathcal{L} lo spazio vettoriale costituito dagli elementi H di \mathcal{P}_b per i quali valga la desiderata proprietà, si vede facilmente che \mathcal{L} è uno spazio vettoriale monotono contenente le indicatori degli intervalli stocastici della forma $[0, T]$, con T tempo

d'arresto (non necessariamente limitato): \mathcal{L} coincide allora con \mathcal{P}_b per il teorema delle classi monotone.

Sia H un elemento di \mathcal{P} , e sia (T_k) una successione crescente di tempi d'arresto, con $\sup_k T_k = \infty$, tale che il processo $HI_{[0, T_k]}$ sia limitato. Scegliamo, per ciascun k , un rappresentante Y^k di $f(HI_{[0, T_k]})$. Utilizzando la proprietà appena dimostrata per gli elementi di \mathcal{P}_b (e il principio di ricolloamento (4.1.8)), si vede ch'esiste un elemento Z di \mathcal{R} , tale che, per ciascun k , il processo $Z|^{T_k}$ sia indistinguibile da Y^k (e sia dunque un rappresentante per $f(HI_{[0, T_k]})$).

Inoltre, la classe d'equivalenza in $\tilde{\mathcal{R}}$ determinata da Z non dipende né dalla scelta della successione (T_k) né dalla scelta del rappresentante Y^k . Se allora si denota con $g(H)$ questa classe d'equivalenza, si definisce un'applicazione g da \mathcal{P} a $\tilde{\mathcal{R}}$. Quest'applicazione prolunga f , è lineare e s'annulla su tutti gli elementi evanescenti di \mathcal{P} . Resta da provare ch'essa possiede la proprietà (b) di (4.2.1). Consideriamo a questo scopo una successione (H^n) di elementi di \mathcal{P} convergente verso 0 in \mathcal{P} . Esiste allora un elemento K di \mathcal{P} tale che si abbia

$$(4.2.13) \quad \sup_n |H^n| \leq K$$

fuori da un insieme evanescente. A meno di sostituire H^n con $H^n I_{\{|H^n| \leq K\}}$, possiamo supporre che la relazione (4.2.13) sia verificata dappertutto. Scegliamo una successione crescente (T_k) di tempi d'arresto, con $\sup_k T_k = \infty$, tale che ciascuno dei processi $KI_{[0, T_k]}$ sia limitato, e denotiamo con Z^n un rappresentante di $g(H^n)$, cioè un elemento di \mathcal{R} tale che, per ciascun k , $Z^n|^{T_k}$ sia un rappresentante di $f(H^n I_{[0, T_k]})$. Grazie all'ipotesi (b), ciascuna delle successioni $(Z^n|^{T_k})_n$ (con k fissato) converge verso 0 in \mathcal{R} : ne segue che, per (4.1.5), anche la successione (Z^n) converge verso 0 in \mathcal{R} . Questo completa la dimostrazione.

Dato un processo X , appartenente a \mathcal{R} , si denoti con ΔX il processo $X - X_-$. Tale processo viene chiamato processo *dei salti* di X . È del tutto evidente che, affinché X sia un processo con quasi tutte le traiettorie continue (nel senso che il processo è indistinguibile da un processo a traiettorie continue), occorre e basta che ΔX sia un processo evanescente.

(4.2.14) Proposizione. *Siano X una semimartingala, H un elemento di \mathcal{P} , Y un rappresentante di $H \bullet X$. Allora:*

(a) ΔY è indistinguibile da $H \Delta X$.

(b) Di conseguenza, se la semimartingala X ha quasi tutte le traiettorie continue, la stessa cosa accade anche per Y .

Dimostrazione. Si denoti con \mathcal{L} lo spazio vettoriale costituito da tutti gli elementi H di \mathcal{P} tali che, per ogni rappresentante Y di $H \bullet X$, il processo ΔY sia indistinguibile da $H \Delta X$. Si vede facilmente che \mathcal{L} contiene le indicatori degli insiemi della forma $[0, T]$, con T tempo d'arresto. Basta dunque provare che, se un elemento H di \mathcal{P} è involuppo superiore di una successione (H^n) di elementi di \mathcal{L} , esso appartiene a \mathcal{L} . A questo scopo, si scelga un rappresentante Y di $H \bullet X$ e, per ciascun intero

positivo n , un rappresentante Y^n di $H^n \bullet X$. La successione (Y^n) converge allora verso Y in \mathcal{R} . A meno di passare ad una sottosuccessione si potrà supporre che (Y^n) converga verso Y puntualmente, al di fuori di un insieme evanescente. D'altra parte, poiché H^n appartiene a \mathcal{L} , ΔY^n è indistinguibile da $H^n \Delta X$. Si ha dunque, fuori da un insieme evanescente,

$$\Delta Y = \lim_n \Delta Y^n = \lim_n H^n \Delta X = H \Delta X,$$

che prova che H appartiene a \mathcal{L} . La conclusione segue dunque attraverso il teorema delle classi monotone. L'asserzione (b) è immediata conseguenza di (a).

4.3 Approssimazione di un integrale stocastico

(4.3.1) Proposizione. *Siano S, T due tempi d'arresto, con $S \leq T$ e sia V una variabile aleatoria reale, misurabile rispetto a \mathcal{F}_S . Sia, inoltre, H il processo continuo a sinistra definito da*

$$H(t, \omega) = V(\omega) I_{\llbracket S, T \rrbracket}(t, \omega).$$

Allora, il processo H è adattato e, dunque, appartiene a \mathcal{P} ed inoltre, per ogni semimartingala X , l'integrale stocastico $H \bullet X$ ammette come rappresentante il processo

$$(t, \omega) \mapsto V(\omega) [X_{t \wedge T}(\omega) - X_{t \wedge S}(\omega)].$$

Dimostrazione. Senza ledere la generalità, possiamo supporre che V sia positiva. Essa è allora involuppo superiore d'una successione crescente di variabili aleatorie semplici, misurabili rispetto a \mathcal{F}_S . Questo permette di ricondurci al caso in cui V sia la funzione indicatrice di un elemento A di \mathcal{F}_S . In questo caso, il processo H , che si può scrivere come $I_{\llbracket S_A, T_A \rrbracket}$, è chiaramente adattato. Per concludere basta osservare che la classe $H \bullet X$ ammette come rappresentante il processo $X^{|T_A} - X^{|S_A}$, cioè:

$$(t, \omega) \mapsto I_A(\omega) [X_{t \wedge T}(\omega) - X_{t \wedge S}(\omega)].$$

(4.3.2) Proposizione. *Sia H un processo della forma*

$$H(t, \omega) = \sum_{k \geq 1} V_{k-1}(\omega) I_{\llbracket T_{k-1}, T_k \rrbracket}(t, \omega),$$

dove (T_k) è una successione crescente di tempi d'arresto e, per ciascun k , V_k è una variabile aleatoria reale misurabile rispetto a \mathcal{F}_{T_k} . Si supponga che, per ogni elemento (t, ω) di $\mathbb{R}_+ \times \Omega$, l'insieme

$$\{T_k(\omega) \wedge t : k \geq 0\}$$

sia finito. Allora il processo H appartiene a \mathcal{P} e, per ogni semimartingala X , l'integrale stocastico $H \bullet X$ ammette come rappresentante il processo

$$(t, \omega) \mapsto \sum_{k \geq 1} V_{k-1}(\omega) [X_{t \wedge T_k}(\omega) - X_{t \wedge T_{k-1}}(\omega)].$$

Dimostrazione. Per ciascun intero $k \geq 1$, il processo H^k definito come

$$H^k(t, \omega) = V_{k-1}(\omega) I_{\llbracket T_{k-1}, T_k \rrbracket}(t, \omega)$$

appartiene a \mathcal{P} per la proposizione precedente. Inoltre, H coincide con il processo $\sum_{k \geq 1} H^k$. Di conseguenza H è adattato e, per le proprietà della successione (T_k) , esso è anche continuo a sinistra ed è localmente limitato. Dunque, che H appartenga a \mathcal{P} segue dalla proposizione (4.1.11). Per concludere, osservato che, per ogni $n \geq 1$, l'integrale stocastico

$$\left(\sum_{k=1}^n H^k\right) \bullet X = \sum_{k=1}^n (H^k \bullet X)$$

ammette come rappresentante (per la proposizione precedente) il processo

$$(t, \omega) \mapsto \sum_{k=1}^n V_{k-1}(\omega) [X_{t \wedge T_k}(\omega) - X_{t \wedge T_{k-1}}(\omega)],$$

la desiderata proprietà segue dal fatto che la successione $(\sum_{k=1}^n H^k)$ converge puntualmente verso H e che, per ogni n , si ha

$$\left|\sum_{k=1}^n H^k\right| = \sum_{k=1}^n |H^k| \leq \sum_{k=1}^{\infty} |H^k| = |H|.$$

(4.3.3) Teorema. Sia X una semimartingala, sia Y un processo appartenente a \mathcal{R} e sia S, T una coppia di tempi d'arresto, con $S \leq T$. Per ciascun intero positivo n , si scelga una successione crescente (T_k^n) di tempi d'arresto, verificanti le relazioni seguenti:

$$T_0^n = S, \quad \lim_n \sup_k T_k^n = T,$$

e tale che la successione (D_n) definita da

$$D_n = \sup_k (T_k^n - T_{k-1}^n)$$

converga verso 0 puntualmente. Si consideri, per ciascun intero positivo n , il processo Z_n definito da

$$Z^n(t, \omega) = \sum_{k \geq 1} Y_{T_{k-1}^n}(\omega) (X_{t \wedge T_k^n}(\omega) - X_{t \wedge T_{k-1}^n}(\omega))$$

Detta Z una versione dell'integrale stocastico $Y_- \bullet X$, la successione (Z^n) converge allora in \mathcal{R} verso il processo $Z|_T - Z|_S$.

Dimostrazione. Senza ledere la generalità possiamo supporre che X sia una semimartingala nulla in 0. Si ponga, per ogni intero positivo n ,

$$H^n(t, \omega) = \sum_{k \geq 1} Y_{T_{k-1}^n}(\omega) I_{\llbracket T_{k-1}^n, T_k^n \rrbracket}(t, \omega).$$

La successione (H^n) converge puntualmente verso $I_{\llbracket S, T \rrbracket} Y_-$ e, siccome il processo H^n appartiene a \mathcal{P} e vale la relazione $|H^n| \leq (Y^*)_-$, la convergenza ha luogo in \mathcal{P} . Ne risulta (si veda (4.3.2)) che l'integrale stocastico $H^n \bullet X$ ammette come rappresentante il processo Z^n . Di conseguenza, la successione (Z^n) converge in \mathcal{R} verso un rappresentante dell'integrale stocastico

$$(I_{\llbracket S, T \rrbracket} Y_-) \bullet X = I_{\llbracket S, T \rrbracket} \bullet Z,$$

cioè verso il processo $Z^{|T} - Z^{|S}$.

(4.3.4) Corollario. *Nelle stesse ipotesi di (4.3.3), si ponga, per ciascun intero positivo n ,*

$$\Sigma_n = \sum_{k \geq 1} Y_{T_{k-1}^n} (X_{T_k^n} - X_{T_{k-1}^n}).$$

La successione (Σ_n) converge allora in probabilità verso un rappresentante dell'integrale stocastico $\int_S^T Y_- dX$.

Dimostrazione. Per concludere basta osservare che, per (4.3.3), la variabile aleatoria

$$(Z^{|T} - Z^{|S})_T = Z_T - Z_S$$

è limite, nel senso della convergenza in probabilità, della successione (Z_T^n) che coincide con la successione (Σ_n) .

4.4 Un esempio di integrazione stocastica: l'integrale di Stieltjes

Un processo X si dice *di Stieltjes* (o *processo a variazione finita*) se è continuo a destra e tale che il processo W definito da

$$(4.4.1) \quad W_t = \sup_{n \geq 1} \sum_{k=1}^{2^n} |X_{tk2^{-n}} - X_{t(k-1)2^{-n}}|$$

sia finito. In tal caso, il processo W prende il nome di *variazione totale* di X .

È chiaro che un processo di Stieltjes X è regolare e, essendo continuo a destra, esso è *progressivo se e soltanto se è adattato*.

Dato un processo di Stieltjes X , si può considerare, per ciascun elemento ω di Ω , la misura con segno di Radon α_ω su $]0, \infty[$ la cui funzione di ripartizione sia la funzione $X(\cdot, \omega) - X(0, \omega)$, ossia la misura determinata dalla relazione

$$\alpha_\omega([0, t]) = X(t, \omega) - X(0, \omega)$$

per ciascun numero reale positivo t . Ciò è possibile perché le traiettorie sono continue a destra e ammettono, per ogni intervallo limitato e chiuso $[0, t]$ una variazione totale finita (uguale a $W(t, \omega)$).

È chiaro che il processo X è completamente determinato dalla famiglia $(\alpha_\omega)_{\omega \in \Omega}$. Inoltre, le misure α_ω sono positive se e soltanto se le traiettorie di X sono crescenti: in tal caso diremo che il processo X è *crescente*; in tal caso, inoltre, si ha $\alpha_\omega([0, t]) = W_t(\omega)$.

(4.4.2) Esempio. Si consideri il processo “deterministico” X definito da

$$X(t, \omega) = t.$$

È immediato riconoscere che si tratta di un processo di Stieltjes: in effetti, per ciascun numero reale e positivo t , si ha

$$W_t = \sup_n \sum_{k=1}^{2^n} |X_{tk2^{-n}} - X_{t(k-1)2^{-n}}| = t.$$

Si riconosce facilmente inoltre che, per ciascun elemento ω di Ω , la misura di Radon α_ω associata a X coincide con la misura di Borel-Lebesgue su \mathbb{R}_+ .

Osserviamo che, se X è un processo di Stieltjes, tale è la sua variazione totale, ossia il processo W : in effetti, per ciascun elemento ω in Ω , la traiettoria $W(\cdot, \omega)$ è la funzione di ripartizione della misura $|\alpha_\omega| = \alpha_\omega \vee (-\alpha_\omega)$. Da questa semplice osservazione discende il seguente risultato:

(4.4.3) Proposizione. *Lo spazio vettoriale costituito da tutti i processi di Stieltjes è uno spazio di Riesz: per ogni processo di Stieltjes X , l’involuppo superiore dei processi X e $-X$ è il processo W (variazione totale di X , definita da (4.4.1)). Inoltre il processo X è somma di due processi crescenti U e V (detti variazione positiva e variazione negativa di X) definiti dalle relazioni*

$$U - V = X, \quad U + V = W.$$

Sia X un processo di Stieltjes e sia H un processo misurabile. Si supponga che, in ogni intervallo limitato contenuto in \mathbb{R}_+ , la funzione $s \mapsto H(s, \omega)$ sia integrabile rispetto alla misura α_ω . Allora, possiamo porre, per ogni elemento (t, ω) in $\mathbb{R}_+ \times \Omega$,

$$(4.4.4) \quad H \cdot X(t, \omega) = H_0(\omega)X_0(\omega) + \int_0^t H(s, \omega) \alpha_\omega(ds),$$

dove il secondo membro denota l’integrale della funzione $s \mapsto H(s, \omega)$ sull’intervallo $]0, t]$ rispetto alla misura α_ω .

(4.4.5) Proposizione. *Sotto le ipotesi precedenti, il processo $H \cdot X$ definito da (4.4.4) è un processo di Stieltjes. Per di più questo processo è:*

- (a) crescente, se X è crescente e H è positivo;
- (b) progressivo (risp. prevedibile), se X e H sono progressivi (risp. prevedibili);
- (c) identico al processo arrestato $X|_T$, se H coincide con il processo $I_{[0, T]}$, dove T è un tempo d’arresto positivo.

Dimostrazione. La proprietà (c) è immediata. Grazie alla decomposizione data da (4.4.3), per provare le altre proprietà possiamo supporre che X sia crescente e H positivo. In questo caso la funzione $t \mapsto (H \cdot X)(t, \omega) - H_0X_0(\omega)$ è la funzione di

ripartizione di una misura di Radon positiva su $]0, \infty[$ (avente $H(\cdot, \omega)$ come densità rispetto alla misura α_ω). È immediato provare che $H \cdot X$ è un processo di Stieltjes crescente.

Per provare che $H \cdot X$ è un processo progressivo se X e H lo sono, ci si riconduce al caso precedente osservando che, per ciascun numero reale positivo t , il processo arrestato $(H \cdot X)^{|t}$ coincide con il processo $(HI_{\llbracket 0, t \rrbracket}) \cdot (X^{|t})$.

Ancora con un ragionamento “per classi monotone”, si prova che $H \cdot X$ è un processo prevedibile se tali sono H e X : in effetti, poiché la proprietà è vera nel caso in cui H sia l'indicatrice di un intervallo stocastico della forma $\llbracket 0, T \rrbracket$, con T tempo d'arresto positivo, è sufficiente osservare che gli insiemi di questo tipo formano una classe stabile per intersezione finita e generano la tribù prevedibile.

Nelle ipotesi precedenti, il processo $H \cdot X$, definito in (4.4.4), è detto *integrale stocastico “di Stieltjes”* di H rispetto a X .

(4.4.6) Teorema. *Sia X un processo di Stieltjes adattato. Allora X è una semimartingala. Inoltre, per ogni elemento H di \mathcal{P} , l'integrale stocastico “di Stieltjes” di H rispetto a X (cioè il processo $H \cdot X$ definito in (4.4.4)) è un rappresentante di $H \bullet X$.*

Dimostrazione. Consideriamo l'applicazione lineare f da \mathcal{P} a $\tilde{\mathcal{R}}$ che, ad ogni elemento H di \mathcal{P} , associa l'elemento di $\tilde{\mathcal{R}}$ avente $H \cdot X$ come rappresentante. Si tratta di provare che quest'applicazione possiede le proprietà (a) e (b) della definizione (4.2.1). In effetti, la proprietà (a) è banalmente verificata grazie all'affermazione (c) di (4.4.5). Resta da verificare la proprietà (b). Questa, però, è un'immediata conseguenza del teorema di convergenza dominata applicato, per ciascun ω , alla misura di Radon α_ω .

4.5 Un esempio di integrazione stocastica: l'integrale di Itô

Si denoti con \mathcal{M}^2 lo spazio vettoriale costituito da tutte le martingale di quadrato integrabile. Evidentemente si tratta di un sottospazio vettoriale di \mathcal{R} , saturo per la relazione d'indistinguibilità. Si denoti con $\tilde{\mathcal{M}}^2$ lo spazio quoziente di \mathcal{M}^2 rispetto a questa relazione d'equivalenza.

Ricordiamo che ogni elemento di \mathcal{M}^2 è chiuso da una variabile aleatoria appartenente a $\mathcal{L}^2(P)$. Inversamente, per ogni elemento Z di $\mathcal{L}^2(P)$ è possibile trovare una martingala di quadrato integrabile da essa chiusa: basta prendere, a questo scopo, il processo X definito prendendo, per ciascun t , una versione X_t di $P[Z|\mathcal{F}_t]$. Ne segue ch'esiste una corrispondenza biunivoca tra lo spazio $\tilde{\mathcal{M}}^2$ e un sottospazio chiuso \mathcal{V} di $L^2(P)$, la quale permette di trasportare la struttura di spazio di Hilbert indotta

su \mathcal{V} da $L^2(P)$ sull'insieme $\widetilde{\mathcal{M}}^2$. Il prodotto scalare in $\widetilde{\mathcal{M}}^2$ è dunque dato dalla (2.4.2):

$$(X | Y)_{\widetilde{\mathcal{M}}^2} = P[X_\infty Y_\infty],$$

dove X_∞, Y_∞ sono variabili aleatorie che chiudono X, Y rispettivamente.

In particolare, con l'ausilio della *disuguaglianza di Doob* (si veda (2.3.14)) è immediato verificare il risultato seguente.

(4.5.1) Proposizione. *Se una successione di elementi di $\widetilde{\mathcal{M}}^2$ converge verso 0 in $\widetilde{\mathcal{M}}^2$, allora converge verso 0 anche in $\widetilde{\mathcal{R}}$.*

Si denoti con λ la misura di Borel-Lebesgue su $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ e si ponga $\mu = \lambda \otimes P$. È chiaro che si tratta di una misura sullo spazio $\mathbb{R}_+ \times \Omega$.

(4.5.2) Teorema. *Sia B un processo di Wiener su (Ω, \mathcal{A}, P) . Esiste allora un'unica applicazione lineare e continua Φ da $\mathcal{L}^2(\mu)$ a $\widetilde{\mathcal{M}}^2$, tale che, per ciascun tempo d'arresto limitato T , $\Phi(I_{[0, T]})$ ammetta $B|_T$ come rappresentante. Inoltre Φ è un'isometria.*

Allo scopo di provare questo teorema, cominciamo col provare il lemma seguente.

(4.5.3) Lemma. *Siano μ una misura positiva su (E, \mathcal{E}) , p un numero reale positivo maggiore di 1 e \mathcal{H} un sottospazio di Riesz di $\mathcal{L}^p(\mu)$ che generi la tribù \mathcal{E} . Si supponga che esista una successione crescente (g_n) di elementi di \mathcal{H}_+ , tale che $\sup_n g_n = 1$. Allora \mathcal{H} è denso in $\mathcal{L}^p(\mu)$.*

Dimostrazione. Si denoti con $\overline{\mathcal{H}}$ la chiusura di \mathcal{H} in $\mathcal{L}^p(\mu)$. È facile verificare che si tratti di uno spazio di Riesz monotono. Per il teorema delle classi monotone, $\overline{\mathcal{H}}$ contiene allora tutte le funzioni misurabili su (E, \mathcal{E}) comprese tra due elementi di $\overline{\mathcal{H}}$. Sia f un elemento di $\mathcal{L}^p(\mu)$. Allora $\overline{\mathcal{H}}$ deve contenere la successione $(f \wedge (ng_n))$ e quindi anche f essendo limite di questa successione in $\mathcal{L}^p(\mu)$. Ciò basta per concludere che $\overline{\mathcal{H}}$ coincide con $\mathcal{L}^p(\mu)$.

Veniamo ora alla dimostrazione del teorema. Si consideri il sottospazio vettoriale \mathcal{H} di $\mathcal{L}^2(\mu)$ costituito da tutti i processi H della forma

$$(4.5.4) \quad H = \sum_{k=1}^n c_k I_{[T_{k-1}, T_k]},$$

dove (T_k) è una successione crescente e finita di tempi d'arresto limitati e (c_k) una successione finita di numeri reali. È facile riconoscere che si tratta di un elemento di $\mathcal{L}^2(\mu)$: infatti, si ha

$$P \left[\int_0^\infty H_s^2 ds \right] = \sum_{k=1}^n c_k^2 (P[T_k] - P[T_{k-1}]) < \infty.$$

Per il lemma precedente è chiaro che \mathcal{H} è denso in $\mathcal{L}^2(\mu)$. Dunque l'unicità di Φ è evidente. Per provare l'esistenza cominciamo con l'associare, ad ogni elemento H di \mathcal{H} della forma (4.5.4), il processo

$$\Phi(H) = \sum_{k=1}^n c_k (B|^{T_k} - B|^{T_{k-1}}).$$

È facile riconoscere che si tratta di una martingala di quadrato integrabile chiusa dalla variabile aleatoria

$$V = \sum_{k=1}^n c_k (B_{T_k} - B_{T_{k-1}}).$$

Mediante il teorema d'arresto (si veda (2.3.3)) alle martingale $(B_t)_{t \geq 0}$ e $(B_t^2 - t)_{t \geq 0}$ è facile verificare che le variabili aleatorie $B_{T_k} - B_{T_{k-1}}$ sono a due a due non correlate e che si ha $P[(B_{T_k} - B_{T_{k-1}})^2] = P[T_k - T_{k-1}]$. Pertanto, si ha

$$P[V^2] = \sum_{k=1}^n c_k^2 P[(B_{T_k} - B_{T_{k-1}})^2] = \sum_{k=1}^n c_k^2 (P[T_k] - P[T_{k-1}]) = P \left[\int_0^\infty H_s^2 ds \right].$$

Ciò significa che l'applicazione lineare $H \mapsto \Phi(H)$ induce, passando al quoziente, un'isometria da \mathcal{H} a $\widetilde{\mathcal{M}}^2$, la quale può essere prolungata, per densità, ad un'isometria Φ da $\mathcal{L}^2(\mu)$ a $\widetilde{\mathcal{M}}^2$ e ciò completa la dimostrazione.

(4.5.5) Definizione. Nelle ipotesi del teorema (4.5.2), l'isometria Φ da $\mathcal{L}^2(\mu)$ a $\widetilde{\mathcal{M}}^2$ è chiamata *isometria di Itô* associata al processo di Wiener B . Inoltre, per ogni elemento H di $\mathcal{L}^2(\mu)$, la classe $\Phi(H)$ si chiama *l'integrale stocastico di Itô di H* (rispetto a B).

(4.5.6) Teorema. Sia B un processo di Wiener. Allora:

- (a) B è una semimartingala;
- (b) per ogni elemento H di $\mathcal{P} \cap \mathcal{L}^2(\mu)$, l'integrale stocastico $H \bullet B$ coincide con l'integrale stocastico di Itô di H rispetto a B .

Dimostrazione. Si denoti con Φ l'isometria di Itô associata a B e si denoti con (T_k) la successione "deterministica" di tempi d'arresto definita da $T_k = k$. Sia H un elemento di \mathcal{P}_b . È chiaro che, per ciascun k , $I_{[0, T_k]} H$ appartiene a $\mathcal{L}^2(\mu)$; si scelga allora un rappresentante Y^k di $\Phi(I_{[0, T_k]} H)$. Attraverso un ragionamento analogo a quello impiegato in (4.2.9), è facile verificare che $Y^{k+1}|_{T_k}$ è indistinguibile da Y^k . Il principio di ricolloamento (4.1.8) mostra allora ch'esiste un elemento Z di \mathcal{R} tale che, per ciascun k , il processo $Z|^{T_k}$ sia indistinguibile da Y^k , e dunque sia un rappresentante di $\Phi(I_{[0, T_k]} H)$. La classe di equivalenza di $\widetilde{\mathcal{R}}$ della quale Z è un rappresentante non dipende che da H : la denoteremo dunque con $f(H)$. Abbiamo così definito un'applicazione lineare f da \mathcal{P}_b a $\widetilde{\mathcal{R}}$. È chiaro che f verifica le condizioni di (4.2.12): in effetti, la condizione (a) è evidente; per la condizione (b), basta osservare che, se H^n converge verso 0 puntualmente al di fuori di un insieme evanescente, e tale che si abbia $\sup_n |H^n| \leq c$, allora, per ogni t , $\Phi(I_{[0, t]} H^n)$ converge verso 0 in $\widetilde{\mathcal{M}}^2$ e quindi in $\widetilde{\mathcal{R}}$ per (4.5.1). In particolare allora $f(H^n)$ converge verso 0 in $\widetilde{\mathcal{R}}$. Ne segue dunque, dalla proposizione (4.2.12), che B è una semimartingala e che, per ciascun elemento H di $\mathcal{P}_b \cap \mathcal{L}^2(\mu)$, l'integrale stocastico $H \bullet B$ coincide con $\Phi(H)$.

Sia ora H un elemento di $\mathcal{P} \cap \mathcal{L}^2(\mu)$. Ricordiamo che H è limite (in \mathcal{P}) d'una successione (H^n) di elementi di $\mathcal{P}_b \cap \mathcal{L}^2(\mu)$. Questa successione converge verso H anche in $\mathcal{L}^2(\mu)$. Dunque, la successione $(H^n \bullet B)$ converge allo stesso tempo verso

$H \bullet B$ (in $\widetilde{\mathcal{R}}$) e verso $\Phi(H)$ (in $\widetilde{\mathcal{M}}^2$ e quindi, per (4.5.1), in $\widetilde{\mathcal{R}}$). Ne segue che $H \bullet B$ e $\Phi(H)$ coincidono. Il teorema è così provato.

La proposizione che segue riassume le proprietà dell'integrale di Itô.

(4.5.7) Proposizione. *Siano B un processo di Wiener, H un elemento di \mathcal{P} e Y un rappresentante di $H \bullet B$. Allora:*

(a) *Se H appartiene a $\mathcal{L}^2(\mu)$, Y è una martingala di quadrato integrabile, e si ha*

$$(4.5.8) \quad P \left[\left(\int_0^t H_s dB_s \right)^2 \right] = P \left[\int_0^t H_s^2 ds \right].$$

(b) *Per ogni elemento H in $\mathcal{L}^2(\mu)$, Y è una famiglia gaussiana centrata, ed ogni Y_t ha come varianza il numero*

$$P \left[\int_0^t H_s^2 ds \right].$$

(c) *Y ha quasi ogni traiettoria continua.*

Dimostrazione. L'affermazione (a) segue dal fatto che ristretta a $\mathcal{L}^2(\mu)$, l'applicazione $H \mapsto H \bullet B$ coincide con l'isometria di Itô. La formula (4.5.2) segue dunque applicando l'isometria di Itô al processo $\widetilde{H} = I_{[0,t]} H$.

Per quanto riguarda l'affermazione (b): per ogni coppia n, k d'interi positivi, e ogni numero reale positivo t , si ponga $t_k^n = tk2^{-n}$. Senza ledere la generalità si potrà supporre che H sia un elemento di $\mathcal{L}^2(\mu)$ limitato. Nel caso generale basterà considerare una successione crescente (H^n) di elementi limitati di $\mathcal{L}^2(\mu)$ (si potrà prendere ad esempio la successione definita da $H^n = H \wedge n \vee (-n)$) e ragionare su ciascun H^n . Allora la successione di processi (H^n) definita da

$$H^n = \sum_{k \geq 1} H_{t_{k-1}^n} (B^{t_k^n} - B^{t_{k-1}^n})$$

converge verso H in $\mathcal{L}^2(\mu)$ e dunque la successione $(H^n \bullet B)$ converge verso $H \bullet B$ in $\widetilde{\mathcal{M}}^2$ perché Φ è un'isometria e, a fortiori, ogni variabile aleatoria Σ_n (definita in (4.3.4)) converge verso Y_t in \mathcal{L}^2 . Ne segue che Y_t è gaussiana e centrata perché tale è ciascuna delle Σ_n . Per concludere basta dunque osservare che, grazie a (4.5.8) si ha:

$$\text{Var}[Y_t] = P \left[\left(\int_0^t H_s dB_s \right)^2 \right] = P \left[\int_0^t H_s^2 ds \right].$$

L'affermazione (c) segue infine dal fatto che B è un processo a traiettorie continue e dalla proposizione (4.2.14).

(4.5.9) Esempio. Sia B un processo di Wiener, e sia T un numero reale positivo. Vogliamo calcolare l'integrale stocastico di Itô

$$(4.5.10) \quad \int_0^T B_t dB_t.$$

Poiché il processo B ha traiettorie continue, una versione di quest'integrale stocastico si trova attraverso il corollario (4.3.4). A questo scopo si ponga, per ogni coppia n, k d'interi positivi, $T_k^n = Tk2^{-n}$. Dal corollario (4.3.4) segue che una versione di (4.5.10) è data dal limite (in probabilità)

$$\begin{aligned} \lim_n \Sigma_n &= \lim_n \sum_{k \geq 1} B_{T_{k-1}^n} (B_{T_k^n} - B_{T_{k-1}^n}) \\ &= \lim_n \sum_{k \geq 1} \left[\frac{1}{2} (B_{T_k^n}^2 - B_{T_{k-1}^n}^2) - \frac{1}{2} (B_{T_k^n} - B_{T_{k-1}^n})^2 \right] \\ &= \frac{1}{2} B_T^2 - \frac{1}{2} \lim_n \sum_{k \geq 1} (B_{T_k^n} - B_{T_{k-1}^n})^2 \\ &= \frac{1}{2} B_T^2 - \frac{1}{2} T, \end{aligned}$$

dove l'ultima uguaglianza segue dal fatto che il limite converge in \mathcal{L}^2 (e quindi in probabilità) verso T : in effetti, posto,

$$S_n = \sum_{k \geq 1} (B_{T_k^n} - B_{T_{k-1}^n})^2,$$

poiché le variabili aleatorie $(B_{T_k^n} - B_{T_{k-1}^n})^2 - (T_k^n - T_{k-1}^n)$ sono a due a due indipendenti e centrate, si ha

$$\begin{aligned} P[(S_n - T)^2] &= \sum_{k \geq 0} P[(B_{T_k^n} - B_{T_{k-1}^n})^2 - (T_k^n - T_{k-1}^n)]^2 \\ &= \sum_{k \geq 1} (T_k^n - T_{k-1}^n)^2 P \left[\left((B_{T_k^n} - B_{T_{k-1}^n})^2 / (T_k^n - T_{k-1}^n) - 1 \right)^2 \right]. \end{aligned}$$

Osservato poi che, per ogni intero positivo k , $(B_{T_k^n} - B_{T_{k-1}^n})^2 / \sqrt{T_k^n - T_{k-1}^n}$ è una variabile aleatoria gaussiana di legge $\mathcal{N}(0, 1)$, la costante reale

$$C = P \left[\left((B_{T_k^n} - B_{T_{k-1}^n})^2 / (T_k^n - T_{k-1}^n) - 1 \right)^2 \right]$$

non dipende da k , e dunque

$$P[(S_n - T)^2] = C \sum_{k \geq 0} (T_k^n - T_{k-1}^n)^2 = CT2^{-n}$$

converge puntualmente verso 0 al tendere di n all'infinito.

(4.5.11) Esempio. Nelle stesse ipotesi dell'esempio precedente, vogliamo calcolare ora l'integrale stocastico di Itô:

$$(4.5.12) \quad \int_0^T t \, dB_t + \int_0^T B_t \, dt.$$

Dal teorema (4.3.4), con le stesse notazioni di (4.5.9), una versione di questo integrale stocastico è data dal limite in probabilità

$$\begin{aligned} \lim_n \Sigma_n &= \lim_n \sum_{k \geq 1} [T_{k-1}^n (B_{T_k^n} - B_{T_{k-1}^n}) + B_{T_{k-1}^n} (T_k^n - T_{k-1}^n)] \\ &= \lim_n \sum_{k \geq 1} [T_k^n B_{T_k^n} - T_{k-1}^n B_{T_{k-1}^n}] \\ &\quad - \lim_n \sum_{k \geq 1} (T_k^n - T_{k-1}^n) (B_{T_k^n} - B_{T_{k-1}^n}) \\ &= TB_T, \end{aligned}$$

dove l'ultima eguaglianza segue dal fatto che il limite è 0: in effetti, per definizione di T_k^n , si ha

$$\lim_n \sum_{k \geq 1} (T_k^n - T_{k-1}^n) (B_{T_k^n} - B_{T_{k-1}^n}) = \lim_n T 2^{-n} B_T = 0.$$

4.6 Variazione quadratica di una semimartingala

(4.6.1) Definizione. Siano X e Y due semimartingale. Si ponga,

$$(4.6.2) \quad [X, Y] = XY - \int X_- \, dY - \int Y_- \, dX,$$

Si chiama allora *variazione quadratica di X* la classe $[X, X]$.

Si osservi che in realtà la notazione (4.6.2) è ambigua perché XY è un processo mentre $\int X_- \, dY$ e $\int Y_- \, dX$ sono elementi di $\tilde{\mathcal{R}}$. In realtà: in effetti il secondo membro di (4.6.2) va inteso come l'elemento di $\tilde{\mathcal{R}}$ costituito da tutti i processi della forma $XY - U - V$, dove U e V sono rappresentanti degli integrali stocastici $\int X_- \, dY$ e $\int Y_- \, dX$.

La relazione (4.6.2), diviene, per $X = Y$,

$$[X, X] = X^2 - 2 \int X_- \, dX.$$

Poiché l'applicazione $(X, Y) \mapsto [X, Y]$ è un'applicazione bilineare e simmetrica, essa possiede un'*identità di polarizzazione*:

$$[X, Y] = \frac{1}{2} ([X + Y, X + Y] - [X, X] - [Y, Y]).$$

(4.6.3) Esempio. Sia B un processo di Wiener su (Ω, \mathcal{A}, P) . Dall'esempio (4.5.9) segue che un rappresentante di $[B, B]$ è dato dal processo deterministico Z definito come

$$Z(t, \omega) = t.$$

Per convincersene basta osservare che ogni rappresentante di $[B, B]$ è un processo della forma $Z = B^2 - 2U$, dove U è un rappresentante di $\int B dB$; poiché, per quanto visto in (4.5.9), possiamo prendere il processo U definito da $U_t = \frac{1}{2}B_t^2 - \frac{1}{2}t$, e dunque si ha

$$Z_t = B_t^2 - 2\left(\frac{1}{2}B_t^2 - \frac{1}{2}t\right) = t.$$

Inoltre, dall'esempio (4.5.11), segue anche che $[B, Z]$ è la classe costituita da tutti i processi evanescenti. Basta osservare, infatti, che $[B, Z]$ è costituito dai processi della forma $Y = BZ - U$, dove U è un rappresentante di (4.5.12); poiché, per quanto visto in (4.5.11), possiamo prendere il processo U definito da $U_t = tB_t$, si ha $Y_t = 0$.

(4.6.4) Esempio. Sia B un processo di Wiener di dimensione 2: denotiamo con B^1, B^2 le sue componenti. Vogliamo trovare $[B^1, B^2]$. Osserviamo che, essendo B^1, B^2 indipendenti, il processo $W = (B^1 + B^2)/\sqrt{2}$ è anch'esso un processo di Wiener. Dunque $[W, W]$ ha come rappresentante il processo Z (definito in (4.6.3)). Ne segue che $2Z$ è un rappresentante di $[B^1 + B^2, B^1 + B^2]$ e dunque, applicando la formula di polarizzazione, si trae

$$[B^1, B^2] = \frac{1}{2}([B^1 + B^2, B^1 + B^2] - [B^1, B^1] - [B^2, B^2]) = 0.$$

(4.6.5) Proposizione. Siano X, Y due semimartingale, e sia S, T una coppia di tempi d'arresto, con $S \leq T$. Per ciascun intero positivo n , sia (T_k^n) una successione crescente di tempi d'arresto, tale che

$$T_0^n = S, \quad \lim_n \sup_k T_k^n = T,$$

e tale che il processo D , definito da

$$D_n = \sup_k (T_k^n - T_{k-1}^n)$$

converga puntualmente verso 0. Posto Z un rappresentante di $[X, Y]$, il processo

$$(4.6.6) \quad \sum_{k \geq 1} (X|^{T_k^n} - X|^{T_{k-1}^n})(Y|^{T_k^n} - Y|^{T_{k-1}^n})$$

converge in \mathcal{R} verso il processo $Z|^{T} - Z|^{S}$. In particolare, posto $S = 0$ e $T = \infty$, il processo (4.6.6) converge verso Z . Infine $[X, Y]$ ha come rappresentante un processo di Stieltjes adattato, crescente se $Y = X$.

Dimostrazione. Grazie a (4.3.3), $Z^{|T} - Z^{|S}$ è limite in \mathcal{R} di

$$\begin{aligned}
V^n &= (XY)^{|T} - (XY)^{|S} \\
&\quad - \sum_{k \geq 1} X_{T_{k-1}^n} (Y^{|T_k^n} - Y^{|T_{k-1}^n}) \\
&\quad - \sum_{k \geq 1} Y_{T_{k-1}^n} (X^{|T_k^n} - X^{|T_{k-1}^n}) \\
&= \sum_{k \geq 1} [(XY)^{|T_k^n} - (XY)^{|T_{k-1}^n} \\
&\quad - X^{|T_{k-1}^n} (Y^{|T_k^n} - Y^{|T_{k-1}^n}) \\
&\quad - Y^{|T_{k-1}^n} (X^{|T_k^n} - X^{|T_{k-1}^n})] \\
&= \sum_{k \geq 1} (X^{|T_k^n} - X^{|T_{k-1}^n}) (Y^{|T_k^n} - Y^{|T_{k-1}^n}).
\end{aligned}$$

In particolare, quando $Y = X$, è evidente che ogni V^n è crescente e dunque tale è Z ; in generale, grazie alla formula di polarizzazione, Z è differenza di due processi crescenti e continui a destra e quindi è un processo di Stieltjes. Infine, che Z sia adattato segue dal fatto che $[X, Y]$ è un elemento di $\tilde{\mathcal{R}}$.

(4.6.7) Corollario. *Nelle stesse ipotesi della proposizione (4.6.5), si ponga, per ciascun intero positivo n ,*

$$V_n = \sum_{k \geq 1} (X_{T_k^n} - X_{T_{k-1}^n})(Y_{T_k^n} - Y_{T_{k-1}^n}).$$

La successione (V_n) converge allora in probabilità verso $Z_T - Z_S$.

Dimostrazione. Con le notazioni della dimostrazione di (4.6.5), per concludere basta osservare che la variabile aleatoria

$$(Z^{|T} - Z^{|S})_T = Z_T - Z_S$$

è limite, nel senso della convergenza in probabilità, della successione V_n .

(4.6.8) Esempio. Come immediata conseguenza della proposizione precedente segue che se X è un processo di Stieltjes continuo, allora $[X, X]$ coincide con la classe costituita da tutti i processi evanescenti: in effetti, scelta la successione $t_k^n = tk2^{-n}$, da (4.6.7) si ha

$$\begin{aligned}
V_n &= \sum_{k \geq 1} (X_{t_k^n} - X_{t_{k-1}^n})^2 \\
&\leq \sup_k |X_{t_k^n} - X_{t_{k-1}^n}| \sum_{k \geq 1} |X_{t_k^n} - X_{t_{k-1}^n}| \\
&\leq (\sup_k |X_{t_k^n} - X_{t_{k-1}^n}|) W_t,
\end{aligned}$$

dove W coincide con la variazione totale di X . È allora immediato riconoscere che la successione (V_n) converge verso 0 in probabilità. In particolare, da questo segue che, se B è un processo di Wiener, allora esso non può essere un processo di Stieltjes dato che la sua variazione quadratica $[B, B]$ dovrebbe essere la classe dei processi evanescenti.

(4.6.9) Proposizione. Siano X, Y due semimartingale. Allora $[X, Y]$ è una semimartingale e, per ogni elemento H di \mathcal{P} , si ha

$$H \bullet [X, Y] = [H \bullet X, Y].$$

Dimostrazione. Senza ledere la generalità possiamo supporre che entrambe le semimartingale X, Y siano nulle in zero. Si denoti con Z un rappresentante di $[X, Y]$, e si ponga, per ciascun elemento H di \mathcal{P} ,

$$\begin{aligned} f(H) &= (H \bullet X)Y - (H \bullet X)_- \bullet Y - Y_- \bullet (H \bullet X) \\ &= (H \bullet X)Y - (H \bullet X)_- \bullet Y - (Y_- H) \bullet X. \end{aligned}$$

Si tratta di provare che l'applicazione lineare f da \mathcal{P} a $\tilde{\mathcal{R}}$ appena definita coincide con l'integrale stocastico $H \mapsto H \bullet Z$, ossia che valgono le condizioni seguenti:

- (a) Per ogni tempo d'arresto T , $Z^{|T}$ è un rappresentante di $f(I_{[0, T]})$.
- (b) f possiede la proprietà di continuità (4.2.6) (b).

Cominciamo col dimostrare la proprietà (a): si denoti con W un rappresentante di $f(I_{[0, T]}) = [X^{|T}, Y]$, e mostriamo che W è indistinguibile da $Z^{|T}$, cioè che, per ciascun numero reale positivo t , W_t è equivalente modulo P a $Z_{t \wedge T}$. Si ponga $t_k^n = tk2^{-n}$ (per ogni coppia n, k d'interi positivi). Il corollario (4.6.7), applicato alla coppia di semimartingale X, Y e alla coppia di tempi d'arresto $0, t \wedge T$ (con $T_k^n = t_k^n \wedge t \wedge T$), mostra che $Z_{t \wedge T}$ è limite in probabilità della successione (V_n) definita da

$$V_n = \sum_{k=1}^{2^n} (X_{t_k^n \wedge T} - X_{t_{k-1}^n \wedge T})(Y_{t_k^n \wedge T} - Y_{t_{k-1}^n \wedge T}).$$

Lo stesso corollario, applicato alla coppia di semimartingale $X^{|T}, Y$ e alla coppia di tempi d'arresto $0, t$ (con $T_k^n = t_k^n \wedge t$), mostra che W_t è limite in probabilità per la successione (V'_n) definita da

$$V'_n = \sum_{k=1}^{2^n} (X_{t_k^n \wedge T} - X_{t_{k-1}^n \wedge T})(Y_{t_k^n} - Y_{t_{k-1}^n}).$$

Basta allora osservare che si ha $V_n = V'_n$ per ciascun intero positivo n .

Infine, per provare la condizione (b), basta osservare che la proprietà (4.2.6) (b) ha luogo separatamente per ciascuna delle tre applicazioni lineari

$$H \mapsto (H \bullet X)Y, \quad H \mapsto (H \bullet X)_- \bullet Y, \quad H \mapsto (Y_- H) \bullet X.$$

Ora, questo è immediato per la prima e l'ultima di queste applicazioni; per la seconda, basta applicare (4.1.14).

(4.6.10) Corollario. *Nelle ipotesi della proposizione precedente, se H e K appartengono a \mathcal{P} , allora si ha*

$$[H \bullet X, K \bullet Y] = HK \bullet [X, Y].$$

Dimostrazione. Applicando (4.6.9) e la proprietà associativa (4.2.8), si ha

$$[H \bullet X, K \bullet Y] = H \bullet [X, K \bullet Y] = H \bullet (K \bullet [X, Y]) = (HK) \bullet [X, Y].$$

(4.6.11) Corollario. *Siano X, Y due semimartingale. Allora XY è anch'essa una semimartingale. In particolare, lo spazio vettoriale di tutte le semimartingale su (Ω, \mathcal{A}, P) costituisce un'algebra.*

Dimostrazione. La tesi segue infatti da (4.6.9) utilizzando (4.2.8) e (4.6.2).

(4.6.12) Corollario. *Nelle stesse ipotesi di (4.6.5), per ciascun elemento H di \mathcal{R} , detto Z un rappresentante dell'integrale stocastico $H_- \bullet [X, Y]$, il processo*

$$V^n = \sum_{k \geq 1} H_{T_{k-1}^n} (X^{|T_k^n} - X^{|T_{k-1}^n})(Y^{|T_k^n} - Y^{|T_{k-1}^n})$$

converge in \mathcal{R} verso il processo $Z^{|T} - Z^{|S}$.

Dimostrazione. Si osservi che, per le proprietà dell'integrale stocastico,

$$H_- \bullet [X, Y] = H_- \bullet (XY) - H_- \bullet (X_- \bullet Y) - H_- \bullet (Y_- \bullet X) = H_- \bullet (XY) - (HX)_- \bullet Y - (HY)_- \bullet X.$$

Grazie a (4.3.3), $Z^{|T} - Z^{|S}$ è limite in \mathcal{R} di

$$\begin{aligned} V^n &= \sum_{k \geq 1} H_{T_{k-1}^n} ((XY)^{|T_k^n} - (XY)^{|T_{k-1}^n}) - \sum_{k \geq 1} H_{T_{k-1}^n} X_{T_{k-1}^n} (Y^{|T_k^n} - Y^{|T_{k-1}^n}) \\ &\quad - \sum_{k \geq 1} H_{T_{k-1}^n} Y_{T_{k-1}^n} (X^{|T_k^n} - X^{|T_{k-1}^n}) \\ &= \sum_{k \geq 1} H_{T_{k-1}^n} [(XY)^{|T_k^n} - (XY)^{|T_{k-1}^n} - X_{T_{k-1}^n} (Y^{|T_k^n} - Y^{|T_{k-1}^n}) \\ &\quad - Y_{T_{k-1}^n} (X^{|T_k^n} - X^{|T_{k-1}^n})] \\ &= \sum_{k \geq 1} H_{T_{k-1}^n} (X^{|T_k^n} - X^{|T_{k-1}^n})(Y^{|T_k^n} - Y^{|T_{k-1}^n}). \end{aligned}$$

4.7 La formula di Itô

Nel seguito, per semplificare la notazione, se X è un processo appartenente a \mathcal{R} , denoteremo con lo stesso simbolo la classe in $\tilde{\mathcal{R}}$ avente X come rappresentante. Le uguaglianze, dunque, dovranno intendersi in $\tilde{\mathcal{R}}$ piuttosto che in \mathcal{R} .

Se X è un processo di Stieltjes crescente, e se H è un elemento di \mathcal{P} , l'integrale stocastico di H rispetto a X coincide con l'integrale di Lebesgue-Stieltjes $\int H_s dX_s$. In particolare, se f è una funzione di classe \mathcal{C}^1 , è facile verificare che si ha, per ogni coppia s, t di numeri reali positivi, con $s \leq t$,

$$(4.7.1) \quad f(X_t) - f(X_s) = \int_s^t f'(X_u) dX_u.$$

D'altra parte, se B è un processo di Wiener, come abbiamo visto nell'esempio (4.5.9),

$$\int_s^t B_u dB_u = \frac{1}{2}(B_t^2 - B_s^2) - \frac{1}{2}(t - s).$$

Posto $f(x) = x^2/2$, la formula precedente equivale a

$$(4.7.2) \quad f(B_t) - f(B_s) = \int_s^t f'(B_u) dB_u + \frac{1}{2} \int_s^t f''(B_u) du,$$

che sembra non coincidere con la formula valida nel caso di un processo di Stieltjes ma invece, come vedremo, la generalizza. Nel seguito vogliamo dimostrare una formula valida per tutte le semimartingale *continue* (una formula analoga ma più complicata è valida per le semimartingale qualsiasi).

(4.7.3) Teorema (formula di Itô). *Sia X una semimartingala e sia f una funzione reale di classe \mathcal{C}^2 . Allora $f(X)$ è anch'essa una semimartingala e vale l'uguaglianza seguente*

$$(4.7.4) \quad f(X_t) - f(X_s) = \int_s^t f'(X) dX + \frac{1}{2} \int_s^t f''(X) d[X, X]$$

Dimostrazione. Supposta vera la formula (4.7.4), è chiaro che $f(X)$ è una semimartingala perché somma delle due semimartingale $\int f'(X) dX$ e $\frac{1}{2} \int f''(X) d[X, X]$ (si veda (4.2.8)). Basta dunque dimostrare la formula (4.7.4). A questo scopo, innanzitutto, per l'additività dell'integrale stocastico, basta provarla per $s = 0$. Per la formula di Taylor, per ogni coppia x, y di numeri reali, si ha

$$f(y) - f(x) = f'(x)(y - x) + \frac{1}{2} f''(x)(y - x)^2 + R(x, y),$$

dove $|R(x, y)| \leq r(|y - x|)(y - x)^2$, e r è una funzione positiva tale che $\lim_{u \downarrow 0} r(u) = 0$.

Si consideri, per ciascun intero positivo m , il tempo d'arresto

$$S_m(\omega) = \inf\{t : |X_t(\omega)| \geq m\}.$$

Allora, il processo arrestato $X|^{S_m}$ è limitato da m e, se la formula (4.7.4) è valida per ciascuno dei processi $X|^{S_m}$ essa deve valere anche per X . Possiamo dunque supporre che X sia limitato. Fissiamo un numero reale positivo t e consideriamo, per ogni intero positivo n , una successione crescente (T_k^n) di tempi d'arresto tali che

$$T_0^n = 0, \quad \lim_n \sup_k T_k^n = t,$$

e tale che la successione (D_n) , definita da

$$D_n = \sup_k (T_k^n - T_{k-1}^n)$$

converga puntualmente verso 0. Allora

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{k \geq 1} [f(X_{T_k^n}) - f(X_{T_{k-1}^n})] \\ &= \sum_{k \geq 1} [f'(X_{T_{k-1}^n})(X_{T_k^n} - X_{T_{k-1}^n}) + \frac{1}{2} f''(X_{T_{k-1}^n})(X_{T_k^n} - X_{T_{k-1}^n})^2] \\ &\quad + \sum_{k \geq 0} R(X_{T_k^n}, X_{T_{k-1}^n}). \end{aligned}$$

Poiché X è continua, si ha $f'(X)_- = f'(X)$ e anche $f''(X)_- = f''(X)$. Dunque, il primo termine converge in probabilità verso un rappresentante dell'integrale stocastico $\int_0^t f'(X) dX$ per (4.3.4) mentre il secondo termine converge in probabilità verso un rappresentante dell'integrale stocastico $\frac{1}{2} \int_0^t f''(X) d[X, X]$ per (4.6.12). Per concludere basta dunque dimostrare che il terzo termine converge verso 0.

In effetti, esso è maggiorato in modulo da

$$\sup_k r(|X_{T_k^n} - X_{T_{k-1}^n}|) \left(\sum_{k \geq 1} (X_{T_k^n} - X_{T_{k-1}^n})^2 \right);$$

ora, osserviamo che la successione $\sum_{k \geq 1} (X_{T_k^n} - X_{T_{k-1}^n})^2$ converge in probabilità verso un rappresentante di $[X, X]_t$ (si veda (4.6.5)), e d'altronde la successione $\sup_k r(|X_{T_k^n} - X_{T_{k-1}^n}|)$ converge puntualmente verso 0 in virtù del fatto che X ha traiettorie continue e che la successione (D_n) converge puntualmente verso 0. La formula (4.7.4), segue dunque per ogni numero reale positivo t , e ciò conclude la dimostrazione.

(4.7.5) Osservazione. Effettivamente, la formula di Itô fornisce la generalizzazione della formula (4.7.1). Quando la semimartingala continua X è un processo di Stieltjes, infatti, si ha $[X, X] = 0$, e dunque la formula (4.7.4) si riduce a (4.7.1). Analogamente, quando X coincide con un processo di Wiener B , la formula (4.7.4) si riduce a (4.7.2) (si veda (4.6.3)).

4.8 I processi di Itô

Sia assegnato un processo di Wiener B , sullo spazio (Ω, \mathcal{A}, P) , munito della filtrazione \mathcal{F} .

(4.8.1) Definizione. Si chiama un *processo di Itô* un processo X appartenente a \mathcal{R} , tale ch'esistano due processi H, K appartenenti a \mathcal{P} , verificanti la relazione

$$(4.8.2) \quad X = X_0 + \int H_s ds + \int K_s dB_s,$$

dove X_0 è il processo "costante" definito da $X_0(t, \omega) = X(0, \omega)$.

(4.8.3) Lemma. Sia X un processo di Itô. I processi H, K verificanti la relazione (4.8.2) sono allora unici (a meno di processi indistinguibili).

Dimostrazione. Si supponga che esistano due coppie $(H, K), (H', K')$ di elementi di \mathcal{P} , tali che si abbia

$$X - X_0 = \int H_s ds + \int K_s dB_s = \int H'_s ds + \int K'_s dB_s.$$

I due processi

$$Y^1 = \int (H_s - H'_s) ds, \quad Y^2 = \int (K'_s - K_s) dB_s,$$

devono allora essere uguali. Tuttavia, Y^1 è un processo di Stieltjes, quindi tale è anche Y^2 . Ma allora, da (4.6.3), (4.6.8) e (4.6.10), si ha

$$0 = [Y^2, Y^2] = (K' - K)^2 \bullet [B, B] = \int (K'_s - K_s)^2 ds.$$

Dunque, necessariamente K' è indistinguibile da K , e quindi $Y^2 = Y^1 = 0$, cosicché anche H' è indistinguibile da H .

È chiaro (in virtù di (4.2.8)) che un processo di Itô è una semimartingala (continua); inoltre, affinché esso sia di Stieltjes, occorre e basta che il processo K (che compare in (4.8.2)) sia evanescente. In ogni caso, tuttavia, è facile riconoscere (utilizzando (4.6.10), la bilinearità dell'applicazione $(X, Y) \mapsto [X, Y]$, (4.6.8) e (4.6.3)) che vale la formula

$$[X, X] = X_0^2 + \int K_s^2 ds.$$

4.9 Cenno sull'integrale stocastico multidimensionale

Sia X un processo a valori in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Allora, per ogni numero reale e positivo t , si avrà $X_t = [X_t^1, \dots, X_t^d]$, per un'opportuna d -upla di processi reali (X^1, \dots, X^d) . Esprimeremo questo fatto dicendo che X è il *blocco* dei processi X^1, \dots, X^d (e scriveremo semplicemente $X = (X^1, \dots, X^d)$).

Si denota con $\mathcal{P}^{(d)}$ (o semplicemente con \mathcal{P} , quando non vi sia rischio d'ambiguità con le dimensioni) lo spazio vettoriale costituito da tutti i processi prevedibili, a valori in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, localmente limitati. È chiaro che, affinché $H = (H^1, \dots, H^d)$ appartenga a $\mathcal{P}^{(d)}$, occorre e basta che ogni H^i appartenga a $\mathcal{P}^{(1)}$, per ogni indice i . La nozione di convergenza definita su $\mathcal{P}^{(1)}$ si estende dunque, senza nessun cambiamento, allo spazio $\mathcal{P}^{(d)}$.

(4.9.1) Definizione. Sia $Z = (Z^1, \dots, Z^d)$ un processo regolare, continuo a destra e nullo in 0. Si dice che Z è una d -semimartingala, se ciascuna delle Z^i è una semimartingala.

Quando Z è una d -semimartingala, resta definita, a partire dall'integrale stocastico di ciascuna delle Z^i , un'applicazione lineare F da $\mathcal{P}^{(d)}$ a $\tilde{\mathcal{R}}$ nel modo seguente: per ogni elemento H di $\mathcal{P}^{(d)}$, si definisce

$$(4.9.2) \quad F(H) = H^1 \bullet Z^1 + \dots + H^d \bullet Z^d.$$

È chiaro che $F(H)$ godrà di proprietà analoghe a quelle dell'integrale stocastico. In effetti, è immediato provare la proposizione seguente:

(4.9.3) Lemma. F è l'unica applicazione lineare, da $\mathcal{P}^{(d)}$ a $\tilde{\mathcal{R}}$, che verifichi le due proprietà seguenti:

(a) Per ogni processo H appartenente a $\mathcal{P}^{(d)}$, ed ogni tempo d'arresto T , si ha

$$F(HI_{[0,T]}) = F(H)^{|T} \quad e \quad F(\mathbb{1}) = Z,$$

dove $\mathbb{1}$ denota il processo costante $\mathbb{1}(t, \omega) = (1, \dots, 1)$.

(b) (proprietà di continuità) Per ogni successione (H^n) di elementi di $\mathcal{P}^{(d)}$ convergente verso 0 in $\mathcal{P}^{(d)}$, la successione $(F(H^n))$ converge verso 0 in $\tilde{\mathcal{R}}$.

(4.9.4) Definizione. La funzione F definita da (4.9.2), si chiama *integrale stocastico (vettoriale)* associato alla d -semimartingala Z . Il suo valore nel punto H è denotato con $H \bullet Z$ (oppure con $\int H dZ$, o anche $\int H_t dZ_t$ quando s'intenda la classe dei processi nulli in 0) e si chiama *integrale stocastico (vettoriale)* di H rispetto a Z .

La teoria di questo integrale stocastico si riconduce alla teoria studiata nelle sezioni precedenti, senza bisogno di nessuna modifica nelle dimostrazioni. Un'altra estensione, invece, merita di essere citata.

Sia $H = (H^1, \dots, H^n)$, con H^1, \dots, H^n appartenenti a $\mathcal{P}^{(d)}$, e sia Z una d -semimartingala. Si definisce, allora, l'integrale stocastico

$$H \bullet Z = (H^1 \bullet Z, \dots, H^n \bullet Z).$$

Anche questa nozione di integrale stocastico si riconduce all'integrale stocastico "vettoriale" già studiato (basta per questo ragionare "componente per componente").

Con la stessa dimostrazione utilizzata nel caso unidimensionale (teorema (4.7.3)), si può dimostrare la seguente generalizzazione della formula di Itô nel caso multidimensionale:

(4.9.5) Teorema (formula di Itô multidimensionale). Siano X^1, \dots, X^d semimartingale, e sia f una funzione da \mathbb{R}^d a \mathbb{R} di classe \mathcal{C}^2 . Si denoti con X il blocco delle X^i . Allora $f(X)$ è anch'essa una semimartingala e vale l'uguaglianza seguente

$$(4.9.6) \quad f(X_t) - f(X_0) = \int_0^t \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X) dX^i + \frac{1}{2} \int_0^t \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X) d[X^i, X^j].$$

(4.9.7) Esempio. Applicando (4.9.5) alla funzione $(x, y) \mapsto xy$, si ottiene che, se X, Y sono due semimartingale continue, allora il prodotto XY è una semimartingala (continua), e inoltre, la formula (4.9.6) si riduce alla *formula di integrazione per parti*:

$$XY = \int X dY + \int Y dX + [X, Y],$$

che si poteva dedurre direttamente dalla definizione di $[X, Y]$.

(4.9.8) Esempio. Applicando (4.9.5) alla coppia di semimartingale (reali) X, Y , dove Y è il processo definito da $Y(t, \omega) = t$, si ottiene, in virtù di (4.6.3),

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds \\ &+ \int_0^t \frac{\partial f}{\partial x}(s, X_s) dX_s \\ &+ \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s) d[X, X]_s. \end{aligned}$$

In particolare, se X coincide con un processo di Wiener B , la formula si semplifica ulteriormente: si ha infatti

$$f(t, B_t) - f(0, 0) = \int_0^t \left[\frac{\partial f}{\partial t}(s, B_s) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(s, B_s) \right] ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s) dB_s.$$

Equazioni differenziali stocastiche

5.1 Definizione e prime proprietà

Nel seguito supponiamo fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, una filtrazione $(\mathcal{F}_t)_{t \geq 0}$ verificante le condizioni abituali. Supponiamo fissato, inoltre, un processo di Wiener $B = (B^1, \dots, B^d)$ di dimensione d .

(5.1.1) Definizione. Siano assegnate le funzioni reali b_i, σ_{ij} , sullo spazio misurabile $(\mathbb{R}_+ \times \mathbb{R}^n, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^n))$ (dove i, j sono una coppia d'interi, con i compreso tra 1 e n e j compreso tra 1 e d), ed una variabile aleatoria Z (a valori in \mathbb{R}^n) indipendente dalla tribù \mathcal{F}_∞ . Si denotino rispettivamente con b e con σ i blocchi relativi alle b_i e alle σ_{ij} e con \mathcal{U} la tribù generata da Z . Sia inoltre s un fissato numero reale positivo. Si dice che un processo $(X_t)_{t \geq s}$ (a valori in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$) è una *soluzione forte dell'equazione differenziale stocastica*

$$(5.1.2) \quad \begin{cases} dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t, \\ X_s = Z, \end{cases}$$

se, detto $(\mathcal{G}_t)_{t \geq 0}$ l'ingrossamento di $(\mathcal{F}_t)_{t \geq 0}$ definito, per ciascun t , da

$$\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{U},$$

valgono le seguenti proprietà:

- (a) il processo X è adattato alla filtrazione $(\mathcal{G}_t)_{t \geq s}$;
- (b) i processi $b_i(X), \sigma_{ij}(X)$ sono prevedibili (rispetto alla filtrazione $(\mathcal{G}_t)_{t \geq s}$) e localmente limitati (nel senso della definizione (4.1.9));
- (c) il processo X verifica la seguente *equazione integrale stocastica*

$$(5.1.3) \quad X_t^i = Z_i + \int_s^t b_i(u, X_u) du + \sum_{j=1}^d \int_s^t \sigma_{ij}(u, X_u) dB_u^j$$

per ciascun numero reale positivo t superiore o eguale a s .

Tramite la notazione “vettoriale”, la (5.1.3) può essere scritta nella seguente forma contratta

$$X_t = Z + \int_s^t b(u, X_u) du + \int_s^t \sigma(u, X_u) dB_u.$$

È facile riconoscere che, se X è una soluzione di (5.1.2), allora essa è un processo di Itô e, dunque, una semimartingala continua (rispetto alla filtrazione \mathcal{G}). Inoltre esso è per giunta un processo di Stieltjes quando ciascuna delle σ_{ij} coincida con la costante 0. Osserviamo poi che la condizione che Z , e quindi \mathcal{U} , sia indipendente da \mathcal{F}_∞ è necessaria se vogliamo che B continui ad essere un processo di Wiener anche rispetto alla filtrazione (\mathcal{G}_t) .

(5.1.4) Esempio (moto browniano geometrico). Si consideri un processo di Wiener unidimensionale B sullo spazio probabilizzato (Ω, \mathcal{A}, P) , munito della filtrazione (\mathcal{F}_t) . Sia Z una variabile aleatoria positiva e indipendente dalla tribù \mathcal{F}_∞ . L'equazione differenziale stocastica

$$\begin{cases} dX_t = \lambda X_t dt + \mu X_t dB_t, \\ X_0 = Z, \end{cases}$$

ammette allora soluzione, per ogni coppia (λ, μ) di numeri reali positivi. In effetti, basta osservare che questa equazione equivale (grazie alla proprietà associativa dell'integrale) alla relazione

$$(5.1.5) \quad \int_0^t 1/X_t dX_t = \lambda t + \mu B_t,$$

e applicare la formula di Itô alla funzione $f(x) = \log x$. Si ottiene, così, dopo aver osservato che vale l'uguaglianza $[X, X] = Z^2 + \int \mu^2 X_t^2 dt$,

$$\log(X_t/Z) = \int_0^t 1/X_s dX_s - \mu^2 t/2,$$

e di qui, sostituendo nella relazione (5.1.5), si trae

$$\log(X_t/Z) = ((\lambda - \mu^2/2)t + \mu B_t),$$

o, ciò ch'è lo stesso,

$$X_t = Z \exp((\lambda - \mu^2/2)t + \mu B_t).$$

Il processo X così definito si chiama *moto browniano geometrico*.

(5.1.6) Esempio (equazione di Ornstein-Uhlenbeck). Si consideri ancora un processo di Wiener B unidimensionale. Fissato un numero reale x , l'equazione differenziale stocastica

$$\begin{cases} dX_t = -\lambda X_t dt + \mu dB_t \\ X_0 = x \end{cases}$$

ha soluzione per ogni coppia (λ, μ) di numeri reali. Per riconoscerlo basta osservare che, in base alla (4.9.7), l'equazione differenziale equivale alla relazione integrale seguente:

$$e^{\lambda t} X_t = x + \int_0^t e^{\lambda s} dX_s + \int_0^t X_s \lambda e^{\lambda s} ds = x + \int_0^t \mu e^{\lambda s} dB_s.$$

Ne segue

$$X_t = e^{-\lambda t} x + e^{-\lambda t} \int_0^t \mu e^{\lambda s} dB_s.$$

Il processo così costruito si chiama *processo di Ornstein-Uhlenbeck*.

(5.1.7) Definizione. Nelle ipotesi della definizione (5.1.1), si dice che (5.1.2) ha soluzione *fortemente unica* se due arbitrarie soluzioni X e \tilde{X} di (5.1.2) sono processi indistinguibili, o, ciò ch'è lo stesso, se è quasi certo l'evento

$$(5.1.8) \quad C = \bigcap_{t \geq s} \{X_t = \tilde{X}_t\} = \bigcap_{t \geq s, t \in \mathbb{Q}} \{X_t = \tilde{X}_t\}.$$

5.2 Un risultato di esistenza ed unicità

Nelle ipotesi della definizione (5.1.1), è evidente che, se ciascuna delle funzioni σ_{ij} coincide con la costante 0, allora l'equazione (5.1.2) si riduce ad un'equazione differenziale ordinaria del prim'ordine: la nozione di equazione differenziale stocastica si può dunque pensare come a una generalizzazione di quella di equazione differenziale ordinaria. Com'è noto, per quest'ultima classe di equazioni vale il celebre teorema di esistenza ed unicità di Cauchy-Lipschitz. È dunque naturale aspettarsi che una generalizzazione di questo teorema valga per le equazioni differenziali stocastiche.

(5.2.1) Teorema (di esistenza ed unicità). *Nelle stesse ipotesi della definizione (5.1.1) sia T un fissato numero reale positivo. Si supponga che le funzioni b e σ verifichino le condizioni seguenti:*

(a) *Esiste un numero reale e positivo C tale che si abbia, per ogni x in \mathbb{R}^n e ogni numero reale positivo t inferiore o eguale a T ,*

$$|b(t, x)| + |\sigma(t, x)| \leq C(1 + |x|).$$

(b) *Esiste un numero reale positivo D tale che si abbia, per ogni coppia x, y di elementi di \mathbb{R}^n e ogni numero reale positivo t inferiore o eguale a T ,*

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq D|x - y|.$$

Per ogni vettore aleatorio Z in $L^2(P)$, esiste allora un'unica soluzione forte (X_t) dell'equazione differenziale (5.1.2), per ciascun numero reale positivo t compreso tra s e T . Tale soluzione appartiene inoltre allo spazio $L^2(\lambda \otimes P)$, dove λ denota la misura di Borel-Lebesgue su $([s, T], \mathcal{B}([s, T]))$.

Dimostrazione. Senza ledere la generalità possiamo supporre che sia $s = 0$. Per ogni elemento Y di $L^2(\lambda \otimes P)$, e ogni numero reale positivo t , si ponga

$$\Phi(Y)_t = Z + \int_0^t b(u, Y_u) du + \int_0^t \sigma(u, Y_u) dB_u.$$

Si definisce così un'applicazione Φ da $L^2(\lambda \otimes P)$ in $\tilde{\mathcal{R}}$. Affinché (5.1.2) abbia soluzione occorre e basta che l'applicazione Φ sia una contrazione. Incominciamo

col provare che, per ogni elemento Y di $L^2(\lambda \otimes P)$, il processo $\Phi(Y)$ appartiene a $L^2(\lambda \otimes P)$. In effetti, se si denota con $\|\cdot\|$ la norma in $L^2(\lambda \otimes P)$, utilizzando la convessità della funzione $x \mapsto |x|^2$, la proprietà (4.5.8), la disuguaglianza di Schwarz-Hölder, e il teorema di Fubini, si ha

$$\begin{aligned}
\|\Phi(Y)\|^2 &= P \left[\int_0^T |\Phi(Y)_t|^2 dt \right] \\
&\leq 3P \left[T|Z|^2 + \int_0^T \left| \int_0^t b(s, Y_s) ds \right|^2 dt + \int_0^T \left| \int_0^t \sigma(s, Y_s) dB_s \right|^2 dt \right] \\
&\leq 3TP[|Z|^2] + 3 \int_0^T tP \left[\int_0^t |b(s, Y_s)|^2 ds \right] dt + 3 \int_0^T P \left[\int_0^t |\sigma(s, Y_s)|^2 ds \right] dt \\
&\leq 3TP[|Z|^2] + 6C^2(1+T) \int_0^T dt \int_0^t ds [1 + P[|Y_s|^2]] \\
&\leq 3TP[|Z|^2] + 6C^2(1+T)(T^2/2 + T\|Y\|^2).
\end{aligned}$$

Resta da provare che Φ è una contrazione, ossia che è lipschitziana di costante inferiore a 1. A questo scopo, supponiamo fissata una coppia Y, \bar{Y} di elementi di $L^2(\lambda \otimes P)$. Si ha allora

$$\begin{aligned}
\|\Phi(Y) - \Phi(\bar{Y})\|^2 &= \int_0^T P \left[|\Phi(Y)_t - \Phi(\bar{Y})_t|^2 \right] dt \\
&\leq 2 \int_0^T P \left[\left| \int_0^t (b(s, Y_s) - b(s, \bar{Y}_s)) ds \right|^2 \right] dt \\
&\quad + 2 \int_0^T P \left[\left| \int_0^t (\sigma(s, Y_s) - \sigma(s, \bar{Y}_s)) dB_s \right|^2 \right] dt \\
&\leq 2 \int_0^T tP \left[\int_0^t |b(s, Y_s) - b(s, \bar{Y}_s)|^2 ds \right] dt \\
&\quad + 2 \int_0^T P \left[\int_0^t |\sigma(s, Y_s) - \sigma(s, \bar{Y}_s)|^2 ds \right] dt \\
&\leq 2(1+T)D^2 \int_0^T \left(\int_0^t P[|Y_s - \bar{Y}_s|^2] ds \right) dt \\
&\leq 2(1+T)TD^2\|Y - \bar{Y}\|^2.
\end{aligned}$$

Ne segue che Φ è una contrazione quando si sostituisca T con un numero reale positivo T_0 , verificante la relazione

$$0 < T_0 < (4D^2)^{-1} \wedge 1.$$

Fissato un siffatto T_0 , per il teorema delle contrazioni (si veda (A.4.2)) esiste un unico elemento X di $L^2(\lambda \otimes P)$, tale che si abbia

$$\Phi(X) = X,$$

ossia tale che, per ogni numero reale t , compreso tra 0 e T_0 , si abbia

$$X_t = Z + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s.$$

Resta da provare che X può essere estesa a tutto $[0, T]$; si consideri a questo scopo l'equazione differenziale

$$\begin{cases} dY_t = b(Y_t) dt + \sigma(Y_t) dB_t \\ Y_{T_0} = Z_1, \end{cases}$$

dove $Z_1 = X_{T_0}$. Con un ragionamento analogo al precedente si ottiene una soluzione $(Y_t)_{t \in [T_0, 2T_0]}$ definita per t compreso tra T_0 e $2T_0$. Induttivamente si trovano un numero finito $N = [T_0/T] + 1$ intervalli della forma $[hT_0, (h+1)T_0]$, sul quale ha soluzione l'equazione differenziale stocastica

$$(5.2.2) \quad \begin{cases} dY_t = b(Y_t) dt + \sigma(Y_t) dB_t \\ Y_{hT_0} = Z_h, \end{cases}$$

dove $Z_h = X_{hT_0}$. Per concludere basta considerare il processo $(X_t)_{t \in [0, T]}$ che coincide (su $[hT_0, (h+1)T_0]$) con la soluzione dell'equazione differenziale (5.2.2).

(5.2.3) Corollario. *Nelle stesse ipotesi del teorema (5.2.1), si denoti con X^x l'unica soluzione dell'equazione differenziale stocastica*

$$\begin{cases} dX_t^x = b(X_t^x) dt + \sigma(X_t^x) dB_t \\ X_0^x = x. \end{cases}$$

Allora, per ogni elemento ω in Ω , l'applicazione $x \mapsto X_t^x(\omega)$ è continua.

Dimostrazione. Basta osservare che la contrazione Φ_x definita, per ogni numero reale positivo t inferiore o eguale a T , da

$$\Phi_x(Y)_t = x + \int_0^t b(s, Y_s) ds + \int_0^t \sigma(s, Y_s) dB_s$$

verifica le ipotesi del teorema delle contrazioni con parametro (si veda (A.4.3)), la conclusione segue dal fatto che, la soluzione X^x è il punto fisso di Φ_x .

(5.2.4) Esempio. Come corollario del teorema appena dimostrato segue che l'equazione del moto browniano geometrico ammette come unica soluzione il processo definito in (5.1.5). Analogamente, l'equazione di Ornstein-Uhlenbeck ammette come unica soluzione il processo definito in (5.1.8).

(5.2.5) Osservazione. Nelle stesse ipotesi del teorema (5.2.1), si supponga che le ipotesi (a) e (b) valgano per ogni numero reale positivo t (o ciò ch'è lo stesso che T coincida con la costante ∞). Allora è facile riconoscere che il ragionamento fatto nella dimostrazione può essere applicato per trovare una soluzione avente \mathbb{R}_+ come insieme dei tempi (o più in generale avente $[s, \infty[$ come insieme dei tempi).

Nel caso particolare in cui b e σ non dipendano dalla variabile t , cioè quando si ha $\partial b / \partial t = 0$ e $\partial \sigma / \partial t = 0$, allora l'ipotesi (a) è conseguenza dell'ipotesi (b).

Per ogni numero reale positivo s e ogni elemento x di \mathbb{R}^n , si denoti con $(X_t^{s,x})$ l'unica soluzione dell'equazione differenziale stocastica

$$(5.2.6) \quad \begin{cases} dX_t^{s,x} = b(t, X_t^{s,x}) dt + \sigma(t, X_t^{s,x}) dB_t \\ X_s^{s,x} = x, \end{cases}$$

con b e σ verificanti le condizioni (5.2.1)(a) e (5.2.1)(b), e si ponga

$$X(t, s, x, \omega) = X_t^{s,x}(\omega).$$

È immediato riconoscere allora che si ha:

(5.2.7) Corollario. *Data una coppia s_1, s_2 di numeri reali positivi, con $s_1 < s_2$, per ogni numero reale positivo t inferiore o eguale a T , si ha:*

$$X(t, s_1, x, \omega) = X(t, s_2, X(s_2, s_1, x, \omega), \omega).$$

5.3 Unicità debole delle soluzioni

(5.3.1) Definizione. Nelle stesse ipotesi di (5.1.1), si supponga che esista un altro spazio probabilizzato $(\Omega', \mathcal{A}', P')$ e, su di esso, una filtrazione $(\mathcal{F}'_t)_{t \geq 0}$ ed un processo di Wiener B' in modo tale che abbia soluzione l'equazione differenziale stocastica

$$\begin{cases} dX'_t = b(X'_t) dt + \sigma(X'_t) dB'_t \\ X'_k = Z', \end{cases}$$

dove Z' è una qualunque variabile aleatoria a valori in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, isonoma a Z (cioè tale che si abbia $Z(P) = Z'(P')$). Si dice allora che (5.1.2) ha *soluzione debole*.

È chiaro che una soluzione forte è, in particolare, una soluzione debole. In generale però si può dimostrare (si veda a questo proposito il celebre “esempio di Tanaka” in [28]) che esistono equazioni differenziali stocastiche dotate di soluzioni deboli ma non di soluzioni forti. (Chiaramente, per il teorema (5.2.1) i coefficienti b, σ non potranno verificare le condizioni (5.2.1)(a) e (5.2.1)(b) perché in tal caso la soluzione forte è garantita.)

(5.3.2) Definizione. Nelle stesse ipotesi di (5.1.1), si dice che (5.1.2) ha *soluzione debolmente unica* se, comunque prese due arbitrarie soluzioni deboli X, X' (definite su due opportuni spazi probabilizzati), hanno la stessa legge i blocchi $[X_t]_{t \geq k}, [X'_t]_{t \geq k}$.

In generale non è detto che un'equazione differenziale stocastica abbia soluzione debolmente unica. Il fatto importante è però che, l'unicità forte assicura quella debole. Più precisamente:

(5.3.3) Teorema. *Nelle ipotesi del teorema (5.2.1), l'equazione differenziale stocastica (5.1.2) ha soluzione debolmente unica.*

Dimostrazione. Basta osservare che, su ciascuno spazio probabilizzato, la soluzione dell'equazione differenziale è limite (in \mathcal{R}) della successione di processi

$$(5.3.4) \quad X_t^{(0)} = Z, \quad X_t^{(n+1)} = Z + \int_0^t b(s, X_s^{(n)}) ds + \int_0^t \sigma(s, X_s^{(n)}) dB_s.$$

La tesi segue dunque osservando che, per induzione su n , hanno la stessa legge i processi (definiti sui rispettivi spazi probabilizzati) della successione (5.3.4).

Il risultato appena provato permette di stabilire che le soluzioni di un'equazione differenziale stocastica, quando i coefficienti b, σ verificano le proprietà (5.2.1)(a) e (5.2.1)(b), non dipendono dallo spazio probabilizzato scelto: in effetti che si scelga uno spazio probabilizzato piuttosto di un altro, la soluzione forte trovata avrà sempre la stessa legge.

(5.3.5) Osservazione. Mettiamoci nelle ipotesi dell'osservazione (5.2.5), ossia supponiamo che le funzioni b e σ non dipendano dalla variabile t . Si denoti ancora con $(X_t^{s,x})$ la soluzione dell'equazione differenziale stocastica (5.2.6). Si verifica allora facilmente, tramite il cambiamento di variabile $u = s + v$, che si ha

$$\begin{aligned} X_{s+t}^{s,x} &= x + \int_s^{s+t} b(X_u^{s,x}) du + \int_s^{s+t} \sigma(X_u^{s,x}) dB_u \\ &= x + \int_s^{s+t} b(X_{s+v}^{s,x}) dv + \int_s^{s+t} \sigma(X_{s+v}^{s,x}) d\tilde{B}_v, \end{aligned}$$

dove \tilde{B} denota il processo di Wiener definito da $\tilde{B}_t = B_{t+s} - B_s$ (si veda (3.2.7)). Posto allora $Y_t = X_{s+t}^{s,x}$, si trae

$$Y_t = x + \int_0^t b(Y_v) dv + \int_0^t \sigma(Y_v) d\tilde{B}_v,$$

e quindi, per l'unicità debole della soluzione, i blocchi $[Y_t]_{t \geq 0}$ e $[X_{t+s}^{s,x}]_{t \geq 0}$ sono isonomi. Si esprime questo fatto dicendo che X è *omogenea nel tempo*. Inoltre, essendo Y_t misurabile rispetto alla tribù $\tilde{\mathcal{F}}_t$, ne segue che essa è anche indipendente dalla tribù \mathcal{F}_s , dato che tale è la tribù $\tilde{\mathcal{F}}_t$. Ciò basta per concludere che il blocco $[Y_t]_{t \geq 0}$ (e dunque $[X_{t+s}^{s,x}]_{t \geq 0}$) è indipendente dal blocco \mathcal{F}_s .

VI

Processi di diffusione

6.1 Natura markoviana delle soluzioni

Nel seguito supporremo sempre assegnato uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, un fissato processo di Wiener B di dimensione d (rispetto alla propria filtrazione naturale). Si denota poi con $(\mathcal{F}_t)_{t \geq 0}$ l'ingrossamento abituale della filtrazione naturale di B (si veda (1.1.7)).

(6.1.1) Definizione. Assegnate le funzioni reali b_i, σ_{ij} , misurabili su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, supporremo sempre che i rispettivi blocchi b, σ verifichino le condizioni (a) e (b) del teorema (5.2.1). Per ogni elemento x in \mathbb{R}^n , si chiama *diffusione di istante iniziale x* l'unica soluzione $(X_t^x)_{t \geq 0}$ dell'equazione differenziale stocastica

$$(6.1.2) \quad \begin{cases} dX_t^x = b(X_t^x) dt + \sigma(X_t^x) dB_t, \\ X_0^x = x. \end{cases}$$

In queste ipotesi, b si chiama *drift* e σ si chiama *coefficiente di diffusione*. Si chiama poi *diffusione* la funzione che ad ogni elemento x di \mathbb{R}^n associa la diffusione di istante iniziale x , cioè il processo $(X_t^x)_{t \geq 0}$. Si denota infine con X^x il blocco $[X_t^x]_{t \geq 0}$.

Siccome i coefficienti b e σ non dipendono dalla variabile t , ogni diffusione è in particolare *omogenea nel tempo*.

(6.1.3) Notazioni. In modo analogo a quanto fatto in (3.2.8), si denoti con W lo spazio $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^n)$, e con $[\xi_t]_{t \geq 0}$ il processo canonico su W , definito da $\xi_t(\omega) = \omega(t)$, per ogni numero reale positivo t , e ogni elemento ω in W . Si ponga, poi,

$$\mathcal{W} = \mathcal{T}(\xi_t : t \geq 0).$$

Ogni diffusione X^x si può pensare come una variabile aleatoria a valori nello spazio misurabile (W, \mathcal{W}) , identificandola con il blocco $[X_t^x]_{t \geq 0}$. Pertanto la sua legge sarà una misura di probabilità su (W, \mathcal{W}) . La si denoti con Q_x . È facile

riconoscere che la famiglia $(Q_x)_{x \in \mathbb{R}^n}$ forma un nucleo markoviano da $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ a (W, \mathcal{W}) , il quale non dipende né dallo spazio probabilizzato (Ω, \mathcal{A}, P) sul quale è definita la diffusione X , né dal processo di Wiener B . In effetti, basta osservare che, per (5.2.3), l'applicazione

$$x \mapsto \int f \, dQ_x = P[f(X^x)]$$

è continua per ogni funzione reale f continua, e quindi misurabile su (W, \mathcal{W}) . Inoltre, dall'unicità debole della soluzione dell'equazione differenziale stocastica (6.1.2), segue che Q_x non dipende né da (Ω, \mathcal{A}, P) , né da B .

Il nucleo markoviano $(Q_x)_{x \in \mathbb{R}^n}$ si denota con Q e si chiama *realizzazione canonica* della diffusione X .

(6.1.4) Notazione. Per ogni numero reale positivo t , ogni elemento x di \mathbb{R}^n , e ogni elemento A di $\mathcal{B}(\mathbb{R}^n)$, si ponga

$$(6.1.5) \quad P_t(x, A) = Q_x\{\xi_t \in A\} = P\{X_t^x \in A\}.$$

È chiaro che, per ogni t , la famiglia di misure di probabilità $(P_t(x, \cdot))_{x \in \mathbb{R}^n}$ forma un nucleo markoviano su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Si riconosce facilmente che la trasformata di una funzione g , limitata e misurabile su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, tramite tale nucleo è data da

$$(6.1.6) \quad P_t g(x) = Q_x[g(\xi_t)] = P[g(X_t^x)].$$

(6.1.7) Osservazione. Sia φ una funzione misurabile da uno spazio misurabile (E, \mathcal{E}) in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Si denoti con $\hat{\varphi}$ il nucleo, da (E, \mathcal{E}) a $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ definito da

$$\hat{\varphi}f = f \circ \varphi \quad \text{per ogni funzione } f \text{ limitata e misurabile su } (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)).$$

Con questa notazione è chiaro che, per ogni numero reale positivo t , il nucleo markoviano P_t coincide con $Q\hat{\xi}_t$ (si veda (A.3.10)). In effetti, per ogni elemento x di \mathbb{R}^n e ogni elemento A di $\mathcal{B}(\mathbb{R}^n)$, si ha (per (A.3.7))

$$Q\hat{\xi}_t I_A(x) = \langle \varepsilon_x, Q\hat{\xi}_t I_A \rangle = \langle Q_x, I_A \circ \xi_t \rangle = Q_x\{\xi_t \in A\}.$$

D'altra parte, per ogni tempo d'arresto T , sullo spazio probabilizzabile (W, \mathcal{W}) , rispetto alla filtrazione naturale del processo canonico, il nucleo $Q\hat{\xi}_T$ è ben definito e dunque si può definire (senza rischio d'ambiguità) il nucleo

$$P_T = Q\hat{\xi}_T.$$

Denotiamo con X un'arbitraria diffusione. Fissato un numero reale positivo s , si denoti con Θ_s il blocco $[X_{t+s}]_{t \geq 0}$, pensato come variabile aleatoria su (Ω, \mathcal{A}, P) , a valori in (W, \mathcal{W}) . Si denoti poi con Q_{X_s} il nucleo markoviano, da (Ω, \mathcal{A}) a (W, \mathcal{W}) , definito da

$$Q_{X_s}(\omega, \cdot) = Q_{X_s(\omega)} \quad \text{per ogni } \omega \in \Omega.$$

(6.1.8) Teorema (Proprietà di Markov). *Per numero reale positivo s , e ogni funzione f , limitata e misurabile su (W, \mathcal{W}) , si ha*

$$P[f(\Theta_s) | \mathcal{F}_s] = Q_{X_s} f.$$

Dimostrazione. Siano s, t una coppia di numeri reali positivi, e sia f una funzione reale, limitata e misurabile su (W, \mathcal{W}) . Si denoti con $(X_t^{s,x})$ l'unica soluzione dell'equazione differenziale stocastica

$$\begin{cases} dX_t = b(X_t) dt + \sigma(X_t) dB_t \\ X_s = x, \end{cases}$$

e si ponga $\Phi_s(x) = [X_{t+s}^{s,x}]_{t \geq 0}$. È chiaro allora, per (5.3.5), che la variabile aleatoria $\Phi_s(x)$ è indipendente dalla tribù \mathcal{F}_s . Inoltre, dal corollario (5.2.7), segue l'uguaglianza

$$\Theta_s = [X_{t+s}]_{t \geq 0} = \Phi_s(X_s).$$

Pertanto, denotata con g la funzione che, ad ogni coppia (x, φ) , costituita da un elemento x di \mathbb{R}^n e una funzione φ misurabile da $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ in (W, \mathcal{W}) , associa il numero reale $f(\varphi(x))$, la variabile aleatoria $f(\Theta_s)$ coincide con $f(\Phi_s(X_s)) = g(X_s, \Phi_s)$, ossia $f(X_{t+s})$ è funzione di X_s e di una variabile aleatoria $\Phi_s(x)$ indipendente dalla tribù \mathcal{F}_s . Ne segue (grazie a (A.2.7)) l'uguaglianza

$$P[f(X_{t+s}) | \mathcal{F}_s] = G(X_s),$$

dove G denota la funzione $x \mapsto P[g(x, \Phi_s)] = P[f(\Phi_s(x))]$. Ora, osservato che, per omogeneità nel tempo, la variabile aleatoria $\Phi_s(x)$ è isonoma al blocco $[X_t^x]_{t \geq 0}$, si trae $P[g(x, \Phi_s)] = P[f([X_t]_{t \geq 0})] = Q_{X_s} f$, e ciò basta per concludere.

Qualche volta, con il termine “proprietà di Markov” s'intende una formulazione più debole di quella presente nel teorema (6.1.8), la quale coinvolge la famiglia (P_t) anziché il nucleo Q_{X_s} . Più precisamente:

(6.1.9) Teorema. *Nelle stesse ipotesi del teorema (6.1.8), per ogni coppia s, t di numeri reali positivi, ed ogni funzione f , limitata e misurabile su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, si ha*

$$P[f(X_{t+s}) | \mathcal{F}_s] = P_t f(X_s).$$

Utilizzando la proprietà di Markov nella forma appena dimostrata, si conclude che la famiglia (P_t) è in realtà un semigrupp.

(6.1.10) Corollario. *La famiglia (P_t) di nuclei markoviani definita in (6.1.5) forma un semigruppulo sullo spazio delle funzioni limitate e misurabili su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ (detto semigruppulo di transizione). Esso è, in particolare, un semigruppulo felleriano (o, ciò ch'è lo stesso, che verifica la proprietà di Feller), nel senso che, per ogni funzione reale f in $\mathcal{C}(\mathbb{R}^n)$, la funzione $P_t f$ appartiene a $\mathcal{C}(\mathbb{R}^n)$ ed inoltre, la famiglia $(P_t f)$ converge puntualmente verso f , per $t \downarrow 0$.*

Dimostrazione. Dobbiamo dimostrare che la famiglia $(P_t)_{t \geq 0}$ verifica l'equazione di Chapman-Kolmogorov, ossia che si ha

$$P_{t+s}f = P_s P_t f \quad \text{per ogni } f \text{ limitata e misurabile su } (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)).$$

A questo scopo, fissiamo una funzione reale f limitata e misurabile su \mathbb{R}^n , dalla proprietà di Markov discende che $P_t f(X_s^x)$ è una versione della speranza condizionale di $P[f(X_{t+s}^x) | \mathcal{F}_s]$ e dunque le due variabili aleatorie $P_t f(X_s^x)$ e $f(X_{t+s}^x)$ hanno la stessa speranza (secondo P). Ne segue

$$(6.1.11) \quad P_{s+t}f(x) = P[f(X_{t+s}^x)] = P[P_t f(X_s^x)].$$

Per concludere basta dunque applicare la (6.1.6) alla funzione $g = P_t f$ ed utilizzare la (6.1.11).

Resta da verificare che (P_t) è felleriano: osserviamo, a questo scopo, che basta provare che, per ogni funzione f continua su \mathbb{R}^n , sono continue le funzioni

$$t \mapsto P_t f(x) = P[f(X_t^x)], \quad x \mapsto P_t f(x) = P[f(X_t^x)].$$

D'altra parte la prima di queste funzioni è continua perché X^x è un processo a traiettorie continue; per quanto riguarda la seconda, basta applicare (5.2.3).

Come semplice applicazione della proprietà di Markov si dimostra la cosiddetta *dicotomia di Blumenthal*:

(6.1.12) Teorema (dicotomia di Blumenthal). *La tribù \mathcal{F}_0 è degenera secondo Q_x , per ogni elemento x in \mathbb{R}^n .*

Dimostrazione. Sia A un elemento di \mathcal{F}_0 . Basta verificare che la funzione I_A è indipendente da se stessa. A questo scopo osserviamo che, per il lemma di misurabilità di Doob (si veda (A.1.16)) esiste una funzione f , limitata e misurabile su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, tale che si abbia $I_A = f(\xi_0)$. Per la proprietà di Markov, si ha allora

$$I_A = Q_x[I_A | \mathcal{F}_0] = Q_x[f(\xi_0) | \mathcal{F}_0] = Q_x[f(\xi_0)] = Q_x[I_A],$$

e ciò basta per concludere che A è un evento degenera.

Le soluzioni delle equazioni differenziali stocastiche godono in realtà di una proprietà più generale della (6.1.8). A questo scopo, per ogni tempo d'arresto T , denotiamo con Q_{X_T} il nucleo markoviano, da (Ω, \mathcal{A}) a (W, \mathcal{W}) definito da

$$Q_{X_T}(\omega, \cdot) = Q_{X_T(\omega)} \quad \text{per ogni } \omega \in \{T < \infty\}.$$

(6.1.13) Teorema (Proprietà forte di Markov). *Nelle ipotesi di (6.1.8), sia T un tempo d'arresto relativo alla filtrazione $(\mathcal{F}_t)_{t \geq 0}$, non equivalente alla costante ∞ . Si denoti con Θ_T il blocco $[X_{T+t}]_{t \geq 0}$, definito su $\{T < \infty\}$. Si ponga*

$$Q = P(\cdot | \{T < \infty\}).$$

Per ogni funzione f , limitata e misurabile su (W, \mathcal{W}) , si ha allora:

$$Q[f(\Theta_T) | \mathcal{F}_T] = Q_{X_T} f.$$

Dimostrazione. Ragionando “per classi monotone”, possiamo ricondurci al caso in cui f appartenga a $\mathcal{C}_b(\mathbb{R}^n)$. Osserviamo innanzitutto che basta dimostrare la proprietà per un tempo d'arresto discreto. Nel caso generale, infatti, basterà considerare una successione crescente (T_n) di tempi d'arresto discreti, ammettenti T come proprio involucro superiore e osservare che, posto $H_n = \{T_n < \infty\}$, dalla relazione

$$P_{H_n}[f(\Theta_{T_n}) | \mathcal{F}_{T_n}] = Q_{X_{T_n}} f$$

la tesi segue passando al limite al tendere di n all'infinito. Supponiamo dunque che T sia discreto. Si tratta di provare che, per ogni evento A in \mathcal{F}_T , si ha

$$\int_A f(\Theta_T) dQ = \int_A Q_{X_T} f dQ.$$

D'altra parte, se $D \cup \{\infty\}$ denota l'insieme numerabile costituito dai valori in T , la tesi equivale a provare che, per ogni evento A in \mathcal{F}_T , e ogni elemento s in D , si ha

$$\int_{A \cap \{T=s\}} f(\Theta_s) dP = \int_{A \cap \{T=s\}} Q_{X_s} f dP,$$

e quest'ultima condizione è verificata, grazie alla proprietà di Markov (6.1.8), perché l'evento $A \cap \{T = s\}$ appartiene alla tribù \mathcal{F}_s .

Anche in questo caso si può enunciare una forma più debole della “proprietà forte di Markov”.

(6.1.14) Teorema. *Nelle stesse ipotesi del teorema (6.1.13), per ogni numero reale e positivo t , si ha*

$$Q[f(X_{t+T}) | \mathcal{F}_t] = P_t f(X_T).$$

(6.1.15) Osservazione. Come abbiamo già osservato, ad ogni diffusione è possibile associare un nucleo markoviano Q , da $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ a (W, \mathcal{W}) . Questo nucleo permette di trasmettere la proprietà di Markov e la proprietà di Markov forte sullo spazio canonico. Più precisamente, sullo spazio canonico (W, \mathcal{W}) , si ponga, per ogni numero reale positivo t ,

$$\mathcal{G}_t = \mathcal{T}(\xi_s : s \leq t)^+.$$

Per ogni tempo d'arresto T , relativo alla filtrazione $(\mathcal{G}_t)_{t \geq 0}$, si denoti con θ_T il processo $[\xi_{T+t}]_{t \geq 0}$ definito su $\{T < \infty\}$. Si denoti poi con Q_{ξ_T} la seguente famiglia di misure di probabilità su \mathcal{W} :

$$(Q_{\xi_T(\omega)})_{\omega \in \{T < \infty\}}.$$

Si ha allora il risultato seguente:

(6.1.16) Teorema (proprietà forte di Markov). *Dato un elemento x di \mathbb{R}^n , e dato un tempo d'arresto T non equivalente alla costante ∞ modulo Q_x , si ponga*

$$Q = Q_x(\cdot | \{T < \infty\}).$$

Per ogni funzione reale f , limitata e misurabile su (W, \mathcal{W}) , si ha allora

$$Q[f(\theta_T) | \mathcal{G}_T] = Q_{\xi_T} f.$$

Come immediato corollario del teorema precedente si ha:

(6.1.17) Corollario. *Sia x un elemento di \mathbb{R}^n , e sia T un tempo d'arresto quasi certamente finito secondo Q_x . Si ha allora, per ogni variabile aleatoria Z limitata su (W, \mathcal{W}) ,*

$$Q_x[Z \circ \theta_T] = Q_x[Q_{\xi_T}[Z]],$$

ossia

$$Q_x[Z \circ \theta_T] = \int \lambda(dy) Q_y[Z],$$

dove λ denota la legge di ξ_T secondo Q_x .

La proprietà di Feller (si veda (6.1.10)) del semigruppato di transizione (P_t) permette di dimostrare il risultato seguente.

(6.1.18) Teorema. Sia A un dominio in \mathbb{R}^n per il quale la funzione

$$T_A(\omega) = \inf\{t > 0 : \xi_t(\omega) \in A^c\}$$

sia un tempo d'arresto limitato. Allora, per ogni numero reale ε strettamente positivo, la funzione

$$T_A^\varepsilon = \varepsilon + T_A \circ \theta_\varepsilon$$

è un tempo d'arresto. Inoltre, si ha

$$T_A^\varepsilon(\omega) = \inf\{t > \varepsilon : \xi_t \in A^c\},$$

e quindi, al tendere di ε verso 0, T_A^ε converge decrescendo verso T_A in modo puntuale.

Dimostrazione. Ci limiteremo a provare che T_A^ε è un tempo d'arresto, ossia che, per ogni istante t maggiore di ε , risulta

$$\{T_A^\varepsilon < t\} \in \mathcal{G}_t.$$

A questo scopo, osserviamo che dalla definizione di T_A^ε discende

$$I_{\{T_A^\varepsilon < t\}} = I_{\{T_A < t-\varepsilon\}} \circ \theta_\varepsilon.$$

D'altra parte, essendo T_A un tempo d'arresto, la funzione $I_{\{T_A < t-\varepsilon\}}$ è misurabile rispetto alla tribù $\mathcal{F}_{t-\varepsilon}$, cioè si può mettere nella forma $f \circ [\xi_s]_{0 \leq s \leq t-\varepsilon}$ (con f funzione misurabile rispetto allo spazio d'arrivo di $[\xi_s]_{0 \leq s \leq t-\varepsilon}$). Dalla relazione precedente discende dunque

$$I_{\{T_A^\varepsilon < t\}} = f \circ [\xi_s]_{0 \leq s \leq t-\varepsilon} \circ \theta_\varepsilon = f \circ [\xi_s]_{\varepsilon \leq s \leq t},$$

e quindi la tesi.

(6.1.19) Corollario. Nelle stesse ipotesi del teorema precedente, valgono le due affermazioni seguenti:

(a) Per ogni numero reale positivo ε , la funzione $x \mapsto Q_x[T_A^\varepsilon]$ è limitata e continua su \mathbb{R}^n .

(b) La funzione $x \mapsto Q_x[T_A]$ è limitata e semicontinua superiormente su \mathbb{R}^n (e quindi continua in ciascun punto in cui si annulla).

Dimostrazione. Basta provare la prima affermazione: la seconda ne discende immediatamente, grazie al fatto che la funzione considerata in (b) è l'involuppo inferiore di una successione decrescente di funzioni del tipo considerato in (a).

Per provare l'affermazione (a), osserviamo che, grazie alla definizione di T_A^ε , e alla proprietà di Markov, si ha

$$Q_x[T_A^\varepsilon] = \varepsilon + Q_x[T_A \circ \theta_\varepsilon] = \varepsilon + Q_x[Q_{\xi_\varepsilon}[T_A]] = \varepsilon + Q_x[f \circ \xi_\varepsilon],$$

dove f denota la funzione così definita: $f(y) = Q_y[T_A]$. Poiché, per ipotesi, T_A è limitato, la conclusione si ottiene applicando la proprietà di Feller (si veda (6.1.10)).

(6.1.20) Esempio. Si consideri la diffusione di equazione

$$dX_t = dB_t,$$

ossia dove ciascuno dei processi (X_t^x) è un processo di Wiener di legge iniziale ε_x . In questo caso possiamo scrivere esplicitamente sia la realizzazione canonica, sia il semigruppato (P_t) in termini della misura di Wiener, che denotiamo con P_w . In effetti, per quanto riguarda la realizzazione canonica, basta osservare che, per ogni misura di probabilità μ sulla tribù $\mathcal{B}(\mathbb{R}^n)$, la trasformata μQ di μ mediante Q è data da $\mu \star P_w$ (si veda a questo proposito (3.2.8)). In particolare, per ogni elemento x in \mathbb{R}^n , si ha $Q_x = \varepsilon_x \star P_w$. Inoltre, per ogni funzione f misurabile e limitata su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, che sia un omomorfismo di gruppi dal gruppo additivo \mathbb{R}^n al gruppo moltiplicativo $\mathbb{R} \setminus \{0\}$, si ha

$$P_t f(x) = P[f(X_t^x)] = f(x)P[f(X_t^x - x)].$$

Ora, osservato che la variabile aleatoria $X_t^x - x$ ha legge $\mathcal{N}(0, t)$, denotato con n_t la densità di probabilità di $\mathcal{N}(0, t)$, si ha

$$P_t f(x) = f(x) \int f(y) \mathcal{N}(0, t)(dy) = \int f(x + y) n_t(y) dy = (f \star n_t)(x).$$

In conclusione, si ha $P_t f = f \star n_t$. Per estendere quest'uguaglianza ad ogni funzione f , limitata e misurabile su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, basta applicare il teorema delle classi monotone.

6.2 Generatore di una diffusione

Sia X una diffusione su \mathbb{R}^n avente b come drift e σ come coefficiente di diffusione. Si denoti con $\mathcal{D}(A)$ l'insieme costituito da tutte le funzioni reali f , definite su \mathbb{R}^n , tali che, per ciascun elemento x di \mathbb{R}^n , esista finito il limite

$$(6.2.1) \quad \lim_{t \downarrow 0} \frac{P_t f(x) - f(x)}{t} = \lim_{t \downarrow 0} \frac{Q_x[f(\xi_t)] - f(x)}{t}.$$

Si denoti con A l'operatore definito su $\mathcal{D}(A)$, che associa, ad ogni elemento f di $\mathcal{D}(A)$, la funzione definita in (6.2.1).

(6.2.2) Definizione. L'operatore A prende il nome di *generatore (infinitesimale)* della diffusione X .

Il seguente teorema caratterizza l'operatore A e permette di legarlo ad un'equazione differenziale. Ricordiamo che si denota con $\mathcal{C}_c^2(\mathbb{R}^n)$ lo spazio vettoriale costituito da tutte le funzioni reali, definite su \mathbb{R}^n , due volte differenziabili con continuità e aventi supporto compatto (anche se ovvio, è forse utile ricordare che per *supporto* di una funzione f s'intende la chiusura (in \mathbb{R}^n) dell'insieme $\{f \neq 0\}$).

(6.2.3) Teorema. Nelle stesse ipotesi di (6.1.1), lo spazio vettoriale $\mathcal{C}_c^2(\mathbb{R}^n)$ è contenuto in $\mathcal{D}(A)$, e si ha

$$Af = \sum_{i=1}^n b_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (\sigma\sigma^*)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \quad \text{per ogni } f \in \mathcal{C}_c^2(\mathbb{R}^n).$$

Dimostrazione. Fissata f in $\mathcal{C}_c^2(\mathbb{R}^n)$, applicando la formula di Itô multidimensionale (si veda (4.9.5)) alla semimartingala X^x , si ha:

$$\begin{aligned} f(X_t^x) - f(x) &= \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(X_s^x) b_i(X_s^x) ds \\ &\quad + \sum_{i=1}^n \sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s^x) \sigma_{ij}(X_s^x) dB_s^j \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s^x) d[X^i, X^j]_s. \end{aligned}$$

Ora, è facile riconoscere (si veda (4.6.3) e (4.6.4)) che si ha

$$[X^i, X^j] = \left[\sum_{\alpha} \sigma_{i\alpha} B^\alpha, \sum_{\beta} \sigma_{j\beta} B^\beta \right] = \sum_{\alpha, \beta} \sigma_{i\alpha} \sigma_{j\beta} [B^\alpha, B^\beta] = (\sigma\sigma^*)_{ij} Z,$$

dove Z è il processo deterministico definito da $Z(t, \omega) = t$. Si ottiene così

$$\begin{aligned} f(X_t^x) - f(x) &= \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(X_s^x) b_i(X_s^x) ds \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s^x) (\sigma\sigma^*)_{ij}(X_s^x) ds + R_t, \end{aligned}$$

ove si sia posto

$$R_t = \sum_{i=1}^n \sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s^x) \sigma_{ij}(X_s^x) dB_s^j.$$

Osserviamo che R_t è una variabile aleatoria centrata, cioè $P[R_t] = 0$: in effetti, poiché f appartiene allo spazio $\mathcal{C}_c^2(\mathbb{R}^n)$, il processo $\frac{\partial f}{\partial x_i}(X^x) \sigma_{ij}(X^x)$ appartiene allo spazio $\mathcal{L}^2(\lambda \otimes P)$, ed il processo Y definito da

$$(6.2.4) \quad Y_t = \int_0^t \frac{\partial f}{\partial x_i}(X_s^x) \sigma_{ij}(X_s^x) dB_s^j$$

è una martingala nulla in 0, e dunque si ha $P[R_t] = 0$. Ne segue, grazie al teorema di Fubini,

$$(6.2.5) \quad \begin{aligned} P[f(X_t^x)] - f(x) &= \sum_{i=1}^n \int_0^t P \left[\frac{\partial f}{\partial x_i}(X_s^x) b_i(X_s^x) \right] ds \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^t P \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(X_s^x) (\sigma \sigma^*)_{ij}(X_s^x) \right] ds, \end{aligned}$$

ossia

$$\begin{aligned} \frac{P_t f(x) - f(x)}{t} &= \sum_{i=1}^n \frac{1}{t} \int_0^t P \left[\frac{\partial f}{\partial x_i}(X_s^x) b_i(X_s^x) \right] ds \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \frac{1}{t} \int_0^t P \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(X_s^x) (\sigma \sigma^*)_{ij}(X_s^x) \right] ds. \end{aligned}$$

La conclusione segue passando al limite per $t \downarrow 0$, e applicando la proprietà di continuità dell'integrale.

(6.2.6) Corollario (formula di Dynkin). *Sia f una funzione appartenente allo spazio $\mathcal{C}_c^2(\mathbb{R}^n)$, e sia T un tempo d'arresto integrabile. Si ha allora*

$$P[f(X_T^x)] = f(x) + P \left[\int_0^T A f(X_s^x) ds \right],$$

Dimostrazione. Basta osservare che la relazione (6.2.5) si estende banalmente al caso in cui si sostituisca t con un arbitrario tempo d'arresto T ed applicare il teorema (6.2.3).

(6.2.7) Osservazione. Se T è il primo istante di uscita di X^x da un insieme limitato U , ed è un tempo d'arresto integrabile (secondo P), allora la formula di Dynkin vale per ogni funzione f appartenente a $\mathcal{C}^2(\mathbb{R}^n)$. Basta osservare, a questo scopo, che se U_ε denota l' ε -involucro di U (ossia l'insieme costituito da tutti i punti di \mathbb{R}^n distanti per meno di ε da U), allora nulla cambia nella formula di Dynkin se si utilizza f oppure una funzione \tilde{f} che coincide con f su U e con 0 su U_ε^c .

(6.2.8) Esempio. Supponiamo che il processo di Wiener B abbia dimensione n . Il generatore associato alla diffusione di equazione

$$dX_t = dB_t$$

è l'operatore

$$A f = \frac{1}{2} \Delta f \quad \text{con } f \in \mathcal{C}_c^2(\mathbb{R}^n),$$

dove Δ denota l'operatore di Laplace definito da $\Delta f = \sum_{i=1}^n \partial^2 f / \partial x_i^2$.

(6.2.9) Esempio. Si consideri l'equazione del moto browniano geometrico (si veda (5.1.4)): si tratta chiaramente di una diffusione. In particolare è immediato riconoscere che il suo generatore è dato dall'operatore differenziale

$$Af(x) = \lambda x f'(x) + \frac{1}{2} \mu^2 x^2 f''(x).$$

Analogamente, nel caso dell'equazione di Ornstein-Uhlenbeck (si veda (5.1.6)), si tratta di una diffusione il cui generatore coincide con l'operatore differenziale seguente:

$$Af(x) = -\lambda x f'(x) + \frac{1}{2} \mu^2 f''(x).$$

(6.2.10) Esempio. Sia B un processo di Wiener di dimensione n . Denotiamo con T_R il primo istante di uscita di B dalla palla $B(0, R)$ di centro 0 e raggio R , ossia T_R è il tempo d'arresto definito da

$$T_R(\omega) = \inf\{t > 0 : |B_t(\omega)| > R\}.$$

Vogliamo calcolare la speranza di T_R . A questo scopo, fissato un numero intero positivo k , applichiamo la formula di Dynkin al tempo d'arresto $T_R \wedge k$. Si ha

$$(6.2.11) \quad P[f(B_{T_R \wedge k})] = f(0) + P \left[\int_0^{T_R \wedge k} \frac{1}{2} \Delta f(B_s) ds \right].$$

Consideriamo una funzione reale f , appartenente a $\mathcal{C}_c^2(\mathbb{R}^n)$, tale che, per ogni x , con $|x| < R$, si abbia $f(x) = |x|^2$. La (6.2.11) diviene allora

$$P[|B_{T_R \wedge k}|^2] = P \left[\int_0^{T_R \wedge k} n ds \right] = nP[T_R \wedge k].$$

Si ha dunque

$$P[T_R \wedge k] = n^{-1} P[|B_{T_R \wedge k}|^2] \leq n^{-1} R^2.$$

Di conseguenza, osservato che, al tendere di k all'infinito, $T_R \wedge k$ converge puntualmente verso T_R , e ricordato che T_R coincide con il primo istante in cui B_t esce da $B(0, R)$, si trae

$$P[T_R] = n^{-1} R^2.$$

Fissiamo una diffusione X e denotiamo con A il suo generatore. Si denoti con $\mathcal{D}(\mathfrak{A})$ l'insieme costituito da tutte le funzioni reali f , definite su \mathbb{R}^n , tal che, per ciascun elemento x di \mathbb{R}^n , esista finito il limite

$$(6.2.12) \quad \lim_{U \in \mathcal{U}_x} \frac{Q_x[f(\xi_{T_U})] - f(x)}{Q_x[T_U]},$$

dove il limite è inteso rispetto al filtro delle sezioni dell'insieme \mathcal{U}_x costituito da tutti gli intorni aperti di x e T_U denota il primo istante di uscita di ξ da U . Si denoti con \mathfrak{A} l'operatore, definito su $\mathcal{D}(\mathfrak{A})$, che associa, ad ogni elemento f di $\mathcal{D}(\mathfrak{A})$, la funzione definita in (6.2.12).

(6.2.13) Definizione. L'operatore \mathfrak{A} prende il nome di *operatore caratteristico* di X .

Sfruttando il fatto che la famiglia filtrata $(T_U)_{U \in \mathcal{U}_x}$ è monotona decrescente, è facile riconoscere che $\mathcal{D}(A)$ è contenuto in $\mathcal{D}(\mathfrak{A})$. È chiaro inoltre che in questo caso si abbia

$$\mathfrak{A}f = Af \quad \text{per ogni } f \in \mathcal{D}(A).$$

Quello che caratterizza l'operatore caratteristico è che esso coincide con l'operatore differenziale definito in (6.2.3) su $\mathcal{C}^2(\mathbb{R}^n)$. Per dimostrarlo occorre il lemma seguente.

(6.2.14) Definizione. Un elemento x di \mathbb{R}^n si dice *assorbente* se

$$Q_x \left(\bigcap_{t \geq 0} \{\xi_t = x\} \right) = 1.$$

È chiaro che x è assorbente se, e soltanto se, il tempo d'arresto $T_{\{x\}}$ (definito come il primo istante d'uscita da $\{x\}$) coincide quasi certamente (secondo Q_x) con la costante ∞ .

(6.2.15) Lemma. Se un elemento x di \mathbb{R}^n non è assorbente (per la diffusione X), allora esiste un intorno aperto U di x , tale che il tempo d'arresto T_U sia integrabile (secondo Q_x).

Dimostrazione. Per ogni numero reale positivo β , si ponga

$$K_\beta = \{y : |x - y| \leq \beta\}.$$

Chiaramente la famiglia filtrata (K_β) è decrescente ed ammette il singoletto $\{x\}$ come involucro inferiore (per $\beta \downarrow 0$). In particolare, per ogni numero reale positivo t , per $\beta \downarrow 0$, si ha

$$P_t(x, K_\beta) = Q_x\{\xi_t \in K_\beta\} \rightarrow P_t(x, \{x\}) = Q_x\{\xi_t = x\}.$$

Ora, poiché x non è un punto assorbente, per ogni numero reale positivo ε è possibile trovare una coppia di numeri reali positivi t, β , tali che si abbia

$$Q_x\{\xi_t \in K_\beta\} = P_t(x, K_\beta) < 1 - \varepsilon.$$

Inoltre, essendo K_β compatto e (P_t) un semigruppello felleriano, l'insieme

$$G = \{y : P_t(y, K_\beta) < 1 - \varepsilon/2\}$$

è aperto. Allora l'insieme $U = G \cap \overset{\circ}{K}_\beta$ è un intorno aperto di x relativamente compatto. Inoltre, per ogni elemento y di U , si ha

$$Q_y\{T_U > t\} \leq P_t(y, K_\beta) < 1 - \varepsilon/2$$

mentre, per ogni elemento y in U^c , si ha

$$Q_y\{T_U > t\} = 0.$$

Ne segue allora

$$\alpha = \sup_{y \in \mathbb{R}^n} Q_y\{T_U > t\} \leq 1 - \varepsilon/2.$$

Osservato (per mezzo della proprietà di Markov) che si ha

$$Q_x\{T_U > t + s\} = \int_{\{T_U > s\}} Q_{\xi_s}\{T_U > t\} dQ_x \leq \alpha Q_x\{T_U > s\},$$

si trae, per induzione,

$$Q_x\{T_U > kt\} \leq \alpha^k.$$

Ne segue, dalle proprietà dell'integrale,

$$Q_x[T_U] \leq \sum_{k \geq 1} tk Q_x\{t(k-1) < T_U \leq kt\} \leq t \sum_{k \geq 1} \alpha^{k-1} = t(1 - \alpha)^{-1}.$$

Pertanto si ottiene la maggiorazione $Q_x[T_U] \leq 2t\varepsilon^{-1}$ e quindi la tesi.

(6.2.16) Teorema. *Nelle ipotesi del teorema (6.2.3), $\mathcal{D}(\mathfrak{A})$ contiene lo spazio vettoriale $\mathcal{C}^2(\mathbb{R}^n)$, e si ha*

$$\mathfrak{A}f = \sum_{i=1}^n b_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (\sigma\sigma^*)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \quad \text{per ogni } f \in \mathcal{C}^2(\mathbb{R}^n).$$

Dimostrazione. Si denoti con L l'operatore differenziale definito sopra. Innanzitutto si supponga che x sia un punto assorbente. Chiaramente allora $\mathfrak{A}f = 0$; d'altra parte, fissato un intorno aperto e limitato U di x , sia \tilde{f} una funzione reale, appartenente a $\mathcal{C}_c^2(\mathbb{R}^n)$, tale che si abbia

$$\tilde{f}(x) = f(x) \quad \text{per ogni } x \in U.$$

Si ha allora grazie al teorema (6.2.3) e all'osservazione precedente,

$$Lf(x) = L\tilde{f}(x) = A\tilde{f}(x) = \mathfrak{A}\tilde{f}(x) = 0.$$

Si supponga adesso che x non sia assorbente: esiste allora, per il lemma (6.2.15), un intorno aperto U di x , tale che il tempo d'arresto T_U sia integrabile. Dalla formula di Dynkin si ottiene

$$\left| \frac{Q_x[f(\xi_{T_U})] - f(x)}{Q_x[T_U]} - Lf(x) \right| = \frac{\left| Q_x \left[\int_0^{T_U} (Lf(\xi_s) - Lf(x)) ds \right] \right|}{Q_x[T_U]} \leq \sup_{y \in U} |Lf(x) - Lf(y)|$$

e di qui, facendo il limite rispetto al filtro delle sezioni dell'insieme \mathcal{U}_x degli intorni di x , si trae $\mathfrak{A}f(x) = Lf(x)$.

6.3 Formula di Feynman-Kač

La formula di Feynman-Kač è la prima applicazione della teoria delle equazioni differenziali stocastiche alla risoluzione di equazioni alle derivate parziali. Precisamente, si ha il risultato seguente:

(6.3.1) Teorema. *Nelle ipotesi della definizione (6.1.1), siano f una funzione reale appartenente a $\mathcal{C}_c^2(\mathbb{R}^n)$, e sia q una funzione reale continua su \mathbb{R}^n . Si supponga che q sia limitata dal basso. Si ponga*

$$(6.3.2) \quad v(t, x) = P \left[f(X_t^x) \exp \left(- \int_0^t q(X_s^x) ds \right) \right].$$

L'applicazione $t \mapsto v(t, x)$ è allora differenziabile e si ha

$$(6.3.3) \quad \begin{cases} \frac{\partial v}{\partial t} = Av - qv, \\ v(0, x) = f(x). \end{cases}$$

Inoltre, se w è una soluzione di (6.3.3) di classe \mathcal{C}^2 , limitata su ogni insieme della forma $K \times \mathbb{R}^n$, con K compatto di \mathbb{R}_+ , allora w è della forma (6.3.2).

Dimostrazione. Per ogni numero reale positivo t , si ponga

$$(6.3.4) \quad Y_t^x = f(X_t^x) \quad Z_t^x = \exp \left(- \int_0^t q(X_s^x) ds \right).$$

È chiaro che Y^x e Z^x sono semimartingale (rispetto alla filtrazione naturale $(\mathcal{F}_t)_{t \geq 0}$ del processo di Wiener B). In particolare Z^x è un processo di Stieltjes, definito dall'equazione differenziale stocastica

$$\begin{cases} dZ_t^x = -Z_t^x q(X_t^x) dt \\ Z_0^x = 1. \end{cases}$$

Ne segue che si ha $[Y^x, Z^x] = 0$ e, di conseguenza, per (4.9.7), si ha

$$Y_t^x Z_t^x = f(x) + \int_0^t Y_s^x dZ_s^x + \int_0^t Z_s^x dY_s^x.$$

Ora, essendo $Y^x Z^x$ un processo di Itô, la funzione $t \mapsto P[Y_t^x Z_t^x] = v(t, x)$ è differenziabile: segue infatti dalla formula di Dynkin che si ha

$$P[Y_t^x Z_t^x] - P[Y_s^x Z_s^x] = \int_s^t P[A(Y_u^x Z_u^x)] du,$$

e di qui, dividendo entrambi i membri per $t - s$, e facendo tendere s verso t , si trae

$$\frac{\partial v}{\partial t}(t, x) = P[A(Y_t^x Z_t^x)].$$

Fissato allora un numero reale positivo r , applicando la proprietà di Markov, si ha

$$P \left[f(X_{t+r}^x) \exp \left(- \int_0^r q(X_{s+r}^x) ds \right) \mid \mathcal{F}_r \right] = g(X_r^x),$$

dove g denota la funzione definita da $g(y) = f(X_t^y) \exp \left(- \int_0^t q(X_s^y) ds \right)$. Passando dunque alle speranze (secondo P), ricordate le notazioni (6.3.4), ed osservato che si ha $P[g(X_r^x)] = P[v(t, X_r^x)]$, si trae

$$P[v(t, X_r^x)] = P \left[Y_{t+r}^x \exp \left(- \int_0^t q(X_{s+r}^x) ds \right) \right].$$

Di qui, osservato che si ha

$$\begin{aligned} P \left[Y_{t+r}^x \exp \left(- \int_0^t q(X_{s+r}^x) ds \right) \right] &= P \left[Y_{t+r}^x \exp \left(- \int_r^{t+r} q(X_u^x) du \right) \right] \\ &= P \left[Y_{t+r}^x Z_{t+r}^x \exp \left(\int_0^r q(X_u^x) du \right) \right] \\ &= P[Y_{t+r}^x Z_{t+r}^x] \\ &\quad + P \left[Y_{t+r}^x \left(\exp \left(\int_r^{t+r} q(X_u^x) du \right) - 1 \right) \right]. \end{aligned}$$

Si ottiene così l'eguaglianza

$$\begin{aligned} r^{-1} (P[v(t, X_t^x) - v(t, x)]) &= r^{-1} (P[Y_{t+r}^x Z_{t+r}^x] - P[Y_t^x Z_t^x]) \\ &\quad + P \left[Y_{t+r}^x \left(\exp \left(\int_r^{t+r} q(X_u^x) du \right) - 1 \right) \right], \end{aligned}$$

e di qui, facendo tendere r verso 0, si trae

$$Av(t, x) = \frac{\partial v}{\partial t}(t, x) + q(x)v(t, x).$$

Resta da provare la seconda parte della tesi. Fissiamo, a questo scopo, una funzione w che verifichi l'equazione (6.3.3) e che sia limitata su ogni insieme della forma $K \times \mathbb{R}^n$, con K compatto di \mathbb{R}_+ . È facile riconoscere allora che w verifica anche l'equazione differenziale

$$\begin{cases} \hat{A}w = -\frac{\partial w}{\partial t} + Aw - qw = 0 & \text{su } \mathbb{R}_+ \times \mathbb{R}^n, \\ w(0, x) = f(x) & \text{per ogni } x \in \mathbb{R}^n. \end{cases}$$

Fissata una terna (s, x, z) in $\mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n$, si denoti con Z la semimartingala (sullo spazio probabilizzato (Ω, \mathcal{A}, P)) definita da

$$Z_t^x = z + \int_0^t q(X_u^x) du.$$

È chiaro allora che il processo tridimensionale H definito da

$$H_t = [s - t, X_t^x, Z_t^x]$$

è una diffusione di generatore

$$A_H \phi = -\frac{\partial \phi}{\partial s} + A\phi + q \frac{\partial \phi}{\partial z} \quad \text{con } \phi \in \mathcal{C}_c^2(\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n).$$

Si denoti con \mathbf{P} la realizzazione canonica di questa diffusione. Applicando la formula di Dynkin alla funzione $\phi(s, x, z) = \exp(-z)w(s, x)$ si ha, per ogni numero reale positivo t , e ogni numero reale strettamente positivo R ,

$$\mathbf{P}_{s,x,z}[\phi(\xi_{t \wedge T_R})] = \phi(s, x, z) + \mathbf{P}_{s,x,z} \left[\int_0^{t \wedge T_R} A_H \phi(\xi_r) dr \right],$$

dove $\mathbf{P}_{s,x,z}$ denota la misura di probabilità $\varepsilon_{(s,x,z)}\mathbf{P}$ e T_R denota il primo istante di uscita da $B(0, R)$ del processo canonico ξ , ossia il tempo d'arresto

$$T_R(\omega) = \inf\{t \geq 0 : |\xi_t(\omega)| > R\}.$$

È chiaro che, per com'è stata scelta ϕ , si ha

$$A_H \phi(s, x, z) = \exp(-z) \left[-\frac{\partial w}{\partial s}(s, x) + Aw(s, x) - q(x)w(s, x) \right] = 0,$$

da cui

$$\begin{aligned} w(s, x) &= \phi(s, x, 0) \\ &= \mathbf{P}_{s,x,0}[\phi(\xi_{t \wedge T_R})] \\ &= P \left[\exp \left(- \int_0^{t \wedge T_R} q(X_r^x) dr \right) w(s - t \wedge T_R, X_{t \wedge T_R}^x) \right]. \end{aligned}$$

Di qui, facendo tendere R all'infinito, poiché w è limitata su ogni insieme della forma $K \times \mathbb{R}^n$, con K compatto di \mathbb{R}_+ , si trae

$$w(s, x) = P \left[\exp \left(- \int_0^t q(X_r^x) dr \right) w(s - t, X_t^x) \right].$$

In particolare, scegliendo $t = s$, si deduce $w(s, x) = v(s, x)$ come desiderato.

VII

Il problema di Dirichlet-Poisson

7.1 Introduzione al problema di Dirichlet-Poisson

Per prima cosa occorre richiamare alcune definizioni della teoria classica delle equazioni differenziali.

(7.1.1) Definizione. Un aperto connesso e relativamente compatto D di \mathbb{R}^n si chiama un *dominio*.

Sia D un fissato dominio in \mathbb{R}^n , e sia L un operatore differenziale lineare del second'ordine su D della forma

$$(7.1.2) \quad L = \sum_{i=1}^n b_i \frac{\partial}{\partial x_i} + \sum_{i,j=1}^n a_{ij} \frac{\partial^2}{\partial x_i \partial x_j},$$

dove b_i, a_{ij} sono funzioni reali definite su D . Supporremo sempre che la matrice $[a_{ij}]$ sia simmetrica.

(7.1.3) Definizione. L'operatore differenziale L si dice *ellittico* in un punto x di D , se la matrice $[a_{ij}(x)]$ ha tutti gli autovalori strettamente positivi, ossia se, denotati con $\lambda(x)$ e $\Lambda(x)$ il più piccolo ed il più grande degli autovalori della matrice $[a_{ij}(x)]$, si ha

$$0 < \lambda(x)|\xi|^2 \leq \langle a(x)\xi, \xi \rangle \leq \Lambda(x)|\xi|^2 \quad \text{per ogni } \xi \in \mathbb{R}^n \setminus \{0\}.$$

In particolare, nel caso in cui sia $\lambda(x) > 0$ per ogni elemento x in D , allora si dice che L è *ellittico in D* .

Si dice, poi, che l'operatore L è *strettamente ellittico*, se esiste una costante reale λ_0 strettamente positiva, tale che si abbia $\lambda > \lambda_0$ su D , ovvero se, per ogni x in D , è verificata la relazione

$$\langle a(x)\xi, \xi \rangle \geq \lambda_0 |\xi|^2 \quad \text{per ogni } \xi \in \mathbb{R}^n \setminus \{0\}.$$

Fissato un operatore ellittico L della forma (7.1.2) vogliamo studiare la risolubilità del problema seguente.

(7.1.4) Problema di Dirichlet-Poisson. *Siano assegnate una funzione reale g continua su D e una funzione reale ϕ continua su ∂D . Trovare (se esiste) una funzione u continua su \overline{D} , e di classe C^2 su D , tale che si abbia*

$$(7.1.5) \quad \begin{cases} Lu = -g & \text{su } D \\ u(x) = \phi(x) & \text{per ogni } x \in \partial D. \end{cases}$$

Per trovare una soluzione di questo problema, fissiamo uno spazio probabilizzato (Ω, \mathcal{A}, P) e, su di esso, un processo di Wiener B di dimensione n (relativo alla propria filtrazione naturale). Si denoti con $(\mathcal{F}_t)_{t \geq 0}$ l'ingrossamento abituale della filtrazione naturale di B . Sia σ una matrice tale che

$$(7.1.6) \quad a = \frac{1}{2} \sigma \sigma^*.$$

Una tale matrice esiste sempre ed è per giunta unica se si richiede che essa sia simmetrica e definita positiva. Denoteremo dunque con σ l'unica matrice simmetrica verificante la relazione (7.1.6). Supporremo poi che b e a (e dunque anche σ) verifichino le condizioni (a) e (b) del teorema (5.2.1). Inoltre è possibile prolungare queste due funzioni a tutto \mathbb{R}^n , in modo tale ch'esse continuino a verificare le condizioni (a) e (b) del teorema (5.2.1). È chiaro allora che l'operatore differenziale L coincide con il generatore della diffusione di equazione

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t.$$

Si denoti con X tale diffusione. Il resto di questo capitolo è dedicato alla ricerca di una formula che leghi la soluzione u del problema (7.1.5) alla diffusione X .

Anche se ovvio, giova forse osservare che lo studio del problema di Dirichlet-Poisson si divide in due parti distinte:

- (a) esistenza della soluzione;
- (b) unicità della soluzione.

La questione dell'esistenza della soluzione è un problema classico, la cui soluzione è stata negli anni ampiamente discussa (si veda ad esempio il testo di D. Gilbarg e N.S. Trudinger [20], oppure E. Giusti [21]). Nel seguito ci interesseremo invece al problema dell'unicità, la cui risoluzione (col metodo stocastico) porta alla costruzione di una formula esplicita per la rappresentazione della soluzione u del problema.

7.2 Teoremi di esistenza ed unicità della soluzione

Nelle stesse ipotesi del problema di Dirichlet-Poisson (7.1.4), e in accordo con le notazioni (6.1.3) e (6.1.4), si denoti con (W, \mathcal{W}) lo spazio canonico, con $(\xi_t)_{t \geq 0}$ il processo canonico, con Q la realizzazione canonica della diffusione X , e con (P_t) il semigruppato di transizione definito da (6.1.5). Fissato un dominio U in D , si denoti poi con T_U il primo istante d'uscita da U del processo (ξ_t) , cioè il tempo d'arresto (sullo spazio canonico) definito da

$$T_U(\omega) = \inf\{t > 0 : \xi_t(\omega) \in U^c\}.$$

(7.2.1) Lemma. *Nelle ipotesi del problema di Dirichlet-Poisson, si considerino le condizioni seguenti:*

(a) *Esiste una funzione reale Φ , definita su \mathbb{R}^n , positiva su D e di classe \mathcal{C}^2 , tale che si abbia*

$$L\Phi \leq -1.$$

(b) *Esistono due costanti reali e positive λ_0, c ed un indice i , compreso tra 1 e n , tale che si abbia*

$$a_{ii} \geq \lambda_0, \quad b_i \geq -c.$$

(c) *La funzione b è limitata su D e l'operatore L è strettamente ellittico. Si ha allora (c) \Rightarrow (b) \Rightarrow (a).*

Dimostrazione. L'implicazione (c) \Rightarrow (b) è ovvia: basta osservare che si ha

$$a_{ii} = \langle ae_i, e_i \rangle \geq \lambda_0 |e_i|^2 = \lambda_0,$$

dove e_i denota il vettore unitario nella direzione d'indice i . Resta da verificare l'implicazione (b) \Rightarrow (a): si osservi, a questo scopo, che, essendo il dominio D relativamente compatto, esiste un numero reale positivo R , e un elemento y di \mathbb{R}^n tale che valga l'inclusione $\overline{D} \subset B(y, R)$. Si ponga, per ogni coppia α, β di numeri reali positivi, e ogni elemento x di \mathbb{R}^n ,

$$\Phi_{\alpha, \beta}(x) = \beta(e^{\alpha R} - e^{\alpha x_i}).$$

È chiaro allora che si ha $\Phi_{\alpha, \beta} > 0$ su D e che $\Phi_{\alpha, \beta}$ è di classe \mathcal{C}^2 su \mathbb{R}^n . Si ha inoltre

$$L\Phi_{\alpha, \beta}(x) = -\beta e^{\alpha x_i} (\alpha^2 a_{ii}(x) + \alpha b_i(x)) \leq -\beta e^{\alpha x_i} (\alpha^2 \lambda_0 - \alpha c) \leq -\alpha \beta e^{-\alpha R} (\alpha \lambda_0 - c).$$

Per concludere basta osservare che, per α, β abbastanza grandi, il secondo membro è inferiore o eguale a -1 .

(7.2.2) Teorema. *Nelle ipotesi di (7.1.4), si supponga che sia verificata la condizione (a) del lemma (7.2.1). Allora il tempo d'arresto T_D è integrabile secondo Q_x , per ogni $x \in D$.*

Dimostrazione. Fissata una coppia s, t di numeri reali positivi, con $s < t$, e fissato un elemento x in D , per la formula di Dynkin (si veda (6.2.6)), si ha

$$Q_x[\Phi(\xi_{t \wedge T_D})] - \Phi(x) = Q_x \left[\int_0^{t \wedge T_D} (L\Phi)(\xi_u) du \right] \leq -Q_x[(t \wedge T_D)].$$

Poiché Φ è positiva su D , si ha $\Phi(x) \geq Q_x[t \wedge T_D]$. Per concludere basta allora passare al limite per $t \uparrow \infty$, ed applicare il lemma di Fatou.

Prima di poter scrivere la formula di rappresentazione della soluzione del problema di Dirichlet-Poisson, occorre stabilire sotto quali condizioni (sull'operatore L e sul dominio D) la soluzione del problema di Dirichlet-Poisson esista veramente. A questo scopo occorre premettere la definizione seguente:

(7.2.3) Definizione. Sia D un dominio in \mathbb{R}^n . Si dice che esso verifica la *condizione della sfera esterna* se, per ogni punto x appartenente a ∂D , esiste una palla $B = B(y, r)$ tale che si abbia $\overline{B} \cap \overline{D} = \{x\}$.

È ben noto il seguente risultato di esistenza ed unicità (per la cui dimostrazione si rimanda a [20], p. 101).

(7.2.4) Teorema (di esistenza ed unicità). *Sia D un dominio e sia L un operatore differenziale strettamente ellittico su D della forma (7.1.2). Si supponga che i coefficienti a_{ij}, b_i appartengano allo spazio $\mathcal{C}^\gamma(D)$ (costituito da tutte le funzioni reali su D , hölderiane di esponente γ) e siano limitati, e che D soddisfi la condizione della sfera esterna (si veda (7.2.3)). Siano poi g una funzione reale e limitata, appartenente a $\mathcal{C}^\gamma(D)$, e ϕ una funzione continua su ∂D . Allora il problema di Dirichlet-Poisson*

$$\begin{cases} Lu = -g & \text{su } D \\ u(x) = \phi(x) & \text{per ogni } x \in \partial D \end{cases}$$

ha un'unica soluzione u appartenente allo spazio $\mathcal{C}(\overline{D}) \cap \mathcal{C}^{2,\gamma}(D)$.

(Anche se ovvio, ricordiamo che si denota con $\mathcal{C}^{k,\gamma}(D)$ il sottospazio vettoriale di $\mathcal{C}^k(D)$, costituito da tutte le funzioni reali, le cui derivate di ordine k siano hölderiane di esponente γ .)

Veniamo adesso all'attesa formula di rappresentazione. Più precisamente, vale il teorema seguente:

(7.2.5) Teorema (formula di rappresentazione). *Nelle stesse ipotesi del teorema (7.2.4), si ha, per ogni elemento x in D ,*

$$(7.2.6) \quad u(x) = Q_x[\phi(\xi_{T_D})] + Q_x \left[\int_0^{T_D} g(\xi_t) dt \right].$$

Dimostrazione. Sia (D_k) una successione crescente di domini, con $D_k \uparrow D$. Si ponga $T_k = T_{D_k}$: è chiaro che T_D è l'involuppo superiore della successione crescente (T_k) , costituita da tempi d'arresto integrabili secondo Q_x , con x in D . Per la formula di Dynkin (si veda (6.2.6)) si ha allora, per ogni elemento x in D_k ,

$$(7.2.7) \quad u(x) = Q_x[u(\xi_{T_k})] - Q_x \left[\int_0^{T_k} (Lu)(\xi_t) dt \right] = Q_x[u(\xi_{T_k})] + Q_x \left[\int_0^{T_k} g(\xi_t) dt \right].$$

Per ipotesi dunque, passando al limite al tendere di k all'infinito, la successione numerica $(u(\xi_{T_k}))$ converge verso $\phi(\xi_{T_D})$. Pertanto si ha

$$\lim_k Q_x[u(\xi_{T_k})] = Q_x[\phi(\xi_{T_D})].$$

Inoltre, è facile riconoscere che si ha

$$\lim_k Q_x \left[\int_0^{T_k} g(\xi_t) dt \right] = Q_x \left[\int_0^{T_D} g(\xi_t) dt \right].$$

In effetti la successione di variabili aleatorie

$$\left(\int_0^{T_k} g(\xi_t) dt \right)$$

converge puntualmente verso la variabile aleatoria $\int_0^{T_D} g(\xi_t) dt$, e dunque, la conclusione si ottiene grazie alla disuguaglianza

$$\left| \int_0^{T_k} g(\xi_t) dt \right| \leq \int_0^{T_D} |g(\xi_t)| dt.$$

Passando dunque al limite in (7.2.7), al tendere di k all'infinito, si trae la conclusione.

Osserviamo che la formula (7.2.6) è interessante da due punti di vista: se sappiamo calcolare le speranze, si ottiene esplicitamente la soluzione del problema di Dirichlet-Poisson; se, viceversa, si conosce la soluzione del problema, si ricavano le speranze delle variabili aleatorie che compaiono in (7.2.6).

Nel resto di questa sezione vogliamo scrivere, in una forma equivalente, la formula di rappresentazione (7.2.6). A questo scopo lavoriamo separatamente nel caso del problema di Dirichlet ($g = 0$) e quello di Poisson ($\phi = 0$).

Nelle ipotesi del teorema (7.2.5), si supponga di avere $g = 0$. La soluzione del problema di Dirichlet si scrive allora nella forma

$$u(x) = Q_x[\phi(\xi_{T_D})].$$

Con le notazioni introdotte in (6.1.7), la funzione u corrisponde alla trasformata della funzione ϕ mediante il nucleo $P_{T_D} = \widehat{Q\xi_{T_D}}$. Tale nucleo si denota anche con H_D e si chiama *nucleo armonico* associato al dominio D : è chiaro che si tratta di un nucleo markoviano da $(\overline{D}, \mathcal{B}(\overline{D}))$ a $(\partial D, \mathcal{B}(\partial D))$, il quale trasforma una funzione f limitata e misurabile su ∂D nella funzione $H_D f$ così definita:

$$H_D f(x) = \int_{\{\xi_0=x, T_D<\infty\}} f(\xi_{T_D}) dQ_x;$$

incidentalmente osserviamo che, per ogni elemento x di \overline{D} , la misura $H_D(x, \cdot)$ coincide con la legge della variabile aleatoria ξ_{T_D} rispetto alla misura di probabilità Q_x , cioè $\xi_{T_D}(Q_x)$. Inoltre, osservato che u coincide con ϕ su ∂D , ne discende il seguente criterio: *affinché una funzione u sia soluzione del problema di Dirichlet, occorre e basta che si abbia*

$$u = H_D \phi.$$

Sempre nelle ipotesi del teorema (7.2.5), si supponga di avere $\phi = 0$. La soluzione del problema di Poisson si scrive allora nella forma

$$u(x) = Q_x \left[\int_0^{T_D} g(\xi_t) dt \right].$$

Per ogni elemento x di D , e ogni elemento A di $\mathcal{B}(D)$, si ponga

$$G_D(x, A) = Q_x \left[\int_0^{T_D} I_A(\xi_t) dt \right].$$

È facile riconoscere che si tratta di un nucleo su $(D, \mathcal{B}(D))$: esso prende il nome di *nucleo di Green* associato al dominio D . In termini di questo nucleo, la funzione u corrisponde alla trasformata della funzione g mediante G_D .

Denotato con G_∞ il nucleo di Green associato al problema con dominio \mathbb{R}^n , si può scrivere questo in termini il semigruppato di transizione (P_t) , utilizzando il teorema di Fubini: si ha cioè

$$G_\infty f(x) = \int_0^\infty P_t f(x) dt.$$

Questa relazione permette anche di scrivere più esplicitamente il nucleo di Green associato ad un dominio D . Si ha infatti, utilizzando la proprietà forte di Markov,

$$\begin{aligned} G_\infty f(x) &= Q_x \left[\int_0^{T_D} f(\xi_t) dt \right] + Q_x \left[\int_{T_D}^\infty f(\xi_t) dt \right] \\ &= G_D f(x) + Q_x \left[\int_0^\infty f(\xi_{t+T_D}) dt \right] \\ &= G_D f(x) + \int_0^\infty P_{t+T_D} f(x) dt \\ &= G_D f(x) + \int_0^\infty P_{T_D} P_t f(x) dt \\ &= G_D f(x) + P_{T_D} G_\infty f(x). \end{aligned}$$

E di qui si trae

$$(7.2.8) \quad G_D f(x) = \int_0^\infty P_t f(x) dt - H_D G_\infty f(x).$$

(7.2.9) Esempio. Supponiamo che L coincida con l'operatore di Laplace: in questo caso il nucleo di Green si può scrivere ancora più esplicitamente. Per quanto osservato in (6.1.20), si ha

$$G_\infty(x, A) = \int_0^\infty (I_A \star n_t)(x) dt = \int_0^\infty dt \int_A n_t(x - y) dy = \int_A 1/(2\pi|x - y|) dy.$$

Inoltre, con un calcolo identico si ottiene

$$\begin{aligned} H_D G_\infty(x, A) &= Q_x [G_\infty f(\xi_{T_D})] \\ &= Q_x \left[\int_0^\infty dt \int_A n_t(\xi_{T_D} - y) dy \right] \\ &= Q_x \left[\int_A 1/(2\pi|\xi_{T_D} - y|) dy \right] \\ &= \int_A dy \int H(x, dz) 1/(2\pi|z - y|). \end{aligned}$$

Ne segue, utilizzando la formula (7.2.8), che il nucleo di Green G_D si scrive come un *nucleo densità*, cioè nella forma

$$G_D(x, A) = \int_A G(x, y) dy,$$

ove si sia posto

$$(7.2.10) \quad G(x, y) = 1/(2\pi|x - y|) - \int H_D(x, dz) 1/(2\pi|z - y|).$$

Osserviamo che la funzione G coincide con la classica *funzione di Green* associata al problema. In questo senso, la (7.2.10) si può considerare come una formula di rappresentazione per la funzione di Green, a patto però di conoscere il nucleo armonico H_D relativo all'operatore di Laplace.

Nel caso generale del problema di Dirichlet-Poisson, abbiamo osservato dunque che il problema

$$\begin{cases} Lu = -g & \text{su } D \\ u(x) = \phi(x) & \text{per ogni } x \in \partial D \end{cases}$$

ha come unica soluzione la funzione

$$u = H_D \phi + G_D g.$$

7.3 Costruzione della soluzione per l'operatore di Laplace

Fissato un dominio D in \mathbb{R}^n , e fissata una funzione reale ϕ , continua su ∂D , consideriamo il problema di Dirichlet

$$(7.3.1) \quad \begin{cases} \Delta u = 0 & \text{su } D, \\ u(x) = \phi(x) & \text{per ogni } x \in \partial D. \end{cases}$$

Senza ricorrere al teorema (7.2.4) costruiremo una soluzione di questo problema. In particolare vedremo a quali ipotesi deve soddisfare il bordo ∂D affinché ci sia la soluzione.

Per quanto osservato precedentemente, la diffusione associata all'operatore di Laplace è il processo di Wiener di dimensione n , pertanto la sua realizzazione canonica è il nucleo $(Q_x)_{x \in \mathbb{R}^n}$ definito da $Q_x = \varepsilon_x \star P$, dove P denota la misura di Wiener. Si denoti con T il primo istante d'uscita del processo canonico $[\xi_t]_{t \geq 0}$ da D . Si ha allora

(7.3.2) Teorema. *Per ogni funzione ϕ , limitata e misurabile su $(\partial D, \mathcal{B}(\partial D))$, la funzione $H_D \phi$ è armonica su D .*

Allo scopo di dimostrare questo risultato occorre premettere il lemma seguente:

(7.3.3) Lemma. *Sia B una palla di centro x . Allora la misura $H_D(x, \cdot)$ è la ripartizione uniforme sulla sfera ∂B .*

Dimostrazione. Senza ledere la generalità possiamo supporre che si abbia $x = 0$. Inoltre denoteremo con S il primo istante d'uscita da B . Allora l'evento

$$(7.3.4) \quad \{\xi_0 = 0, S < \infty\}$$

è quasi certo secondo P e, per ogni elemento ω di questo evento, $\xi_S(\omega)$ è il punto nel quale la traiettoria ω incontra per la prima volta la sfera ∂B . La legge di ξ_S secondo P , ossia $H_D(0, \cdot)$, è una legge su ∂B . Si tratta di provare che essa è invariante rispetto ad ogni rotazione di \mathbb{R}^n intorno all'origine. Fissata una tale rotazione ρ , si denoti con $\tilde{\rho}$ l'applicazione di W in W così definita:

$$\tilde{\rho}(\omega) = \rho \circ \omega.$$

Questa applicazione può essere considerata come un processo, con traiettorie continue, sullo spazio probabilizzato (W, \mathcal{W}, P) . Più precisamente, si tratta di un processo di Wiener uscente dall'origine, ossia la cui legge (sullo spazio (W, \mathcal{W})) è la legge di Wiener. Ne risulta la seguente semplice eguaglianza

$$(7.3.5) \quad P = \tilde{\rho}(P).$$

D'altra parte, per ogni elemento ω dell'evento (7.3.4), l'istante

$$\min \{t \geq 0 : \xi_t(\tilde{\rho}(\omega)) \in \partial B\}$$

coincide con $S(\omega)$ (come facilmente si riconosce osservando che la rotazione ρ conserva le distanze dall'origine). Dunque, se ω (e quindi $\tilde{\rho}(\omega)$) è un elemento dell'evento (7.3.4), allora $\xi_S(\tilde{\rho}(\omega))$ è il punto nel quale la traiettoria $\tilde{\rho}(\omega)$ incontra per la prima volta ∂B . Perciò esso è identico a $\rho(\xi_S(\omega))$:

$$\rho(\xi_S(\omega)) = \xi_S(\tilde{\rho}(\omega)).$$

Da questa eguaglianza, passando alle leggi, si trae

$$\rho(\xi_S(P)) = \xi_S(\tilde{\rho}(P)) = \xi_S(P)$$

(dove l'ultima eguaglianza è dovuta a (7.3.5)). È così provata l'asserita invarianza di $\xi_S(P)$ rispetto alla rotazione ρ .

Veniamo ora alla dimostrazione del teorema (7.3.2). Consideriamo un punto x in D ed una palla aperta B , di centro x , con chiusura contenuta in D . Denotiamo inoltre con S il primo istante d'uscita da B . Allora l'evento

$$\{\xi_0 = x, S < T < \infty\}$$

è quasi certo secondo Q_x . Inoltre, su di esso, ξ_T coincide con $\xi_T \circ \theta_S$. Se dunque si denota con λ la legge di ξ_S secondo Q_x (identica, in virtù del lemma precedente, alla ripartizione uniforme su ∂B) e si applica la proprietà di Markov, si trae

$$\begin{aligned} H_D \phi(x) &= Q_x[\phi \circ \xi_T] \\ &= Q_x[\phi \circ \xi_T \circ \theta_S] \\ &= \int_{\partial B} \lambda(dy) Q_y[\phi \circ \xi_T] \\ &= \int_{\partial B} \lambda(dy) H_D \phi(y). \end{aligned}$$

È così provato che la restrizione di $H_D \phi$ a D possiede la proprietà della media sulle sfere, ossia è armonica.

Osserviamo che l'evento $\{T = 0\}$ appartiene alla tribù \mathcal{G}_0 , la quale, grazie alla dicotomia di Blumenthal, è degenere secondo ciascuna delle misure Q_x . È dunque giustificata la definizione seguente:

(7.3.6) Definizione. Dato un punto x della frontiera di D , si dice che esso è, per l'insieme D , un *punto di uscita regolare* se l'evento $\{T = 0\}$ è quasi certo secondo Q_x . Nell'altro caso, cioè se l'evento $\{T = 0\}$ è trascurabile secondo Q_x , si dice che x è, per l'insieme D , un *punto di uscita irregolare*.

Denoteremo con $\partial_r D$ la parte di ∂D costituita dai punti che sono, per l'insieme D , punti di uscita regolari.

Il teorema che segue mostra l'importanza della precedente definizione.

(7.3.7) Teorema. *Sia ϕ una funzione limitata e misurabile su ∂D . Allora la funzione $H_D\phi$ coincide con ϕ su $\partial_r D$. Inoltre, essa è continua in ciascun punto di $\partial_r D$ nel quale sia continua la funzione ϕ .*

Dimostrazione. La prima affermazione è conseguenza immediata della definizione di $\partial_r D$. Per provare la seconda affermazione, fissiamo un punto z di $\partial_r D$ nel quale la funzione ϕ sia continua, e mostriamo che $H_D\phi(x)$ converge verso $\phi(x)$ al tendere di x verso z . A questo scopo, cominciamo con l'introdurre una notazione: per ogni punto x di \overline{D} , denotiamo con T_x il tempo d'uscita dall'insieme aperto $D - x$. Si ha allora, per ogni ω ,

$$\begin{aligned} T_x(\omega) &= \inf\{t > 0 : \xi_t(\omega) \notin D - x\} \\ &= \inf\{t > 0 : x + \xi_t(\omega) \notin D\} \\ &= T(x + \xi_t(\omega)). \end{aligned}$$

Si può perciò scrivere:

$$(7.3.8) \quad T_x = T \circ (x + \xi), \quad x + \xi_{T_x} = \xi_T \circ (x + \xi).$$

D'altra parte, il processo $x + \xi$ è, secondo P , un processo di Wiener uscente da x , ossia verifica la relazione $(x + \xi)(P) = Q_x$. Le relazioni (7.3.8) implicano dunque

$$T_x(P) = T((x + \xi)(P)) = T(Q_x), \quad (x + \xi_{T_x})(P) = \xi_T((x + \xi)(P)) = \xi_T(Q_x)$$

e, in particolare,

$$P[T_x] = Q_x[T].$$

ed anche

$$(7.3.9) \quad \int_{\{\xi_0=0, T_x<\infty\}} \phi(x + \xi_{T_x}) dP = \int_{\{\xi_0=x, T<\infty\}} \phi(\xi_T) dQ_x = H_D\phi(x).$$

Si vede allora che la funzione $x \mapsto P[T_x]$, essendo identica alla funzione $x \mapsto Q_x[T]$, si annulla sull'insieme $\partial_r D$ ed è quindi continua in ciascun punto di questo insieme (si veda a questo proposito (6.1.19)(b)). In altri termini, l'applicazione $x \mapsto T_x$, considerata come applicazione di \overline{D} in $L^1(P)$, è nulla, e continua, in ciascun punto di $\partial_r D$. Perciò, comunque si assegni una successione di punti di \overline{D} convergente verso z , è possibile estrarne una sottosuccessione (x_k) tale che, su un opportuno evento H quasi certo secondo P , la successione (T_{x_k}) converga puntualmente verso 0. Sull'evento $H \cap \{\xi_0 = 0\}$ (anch'esso quasi certo secondo P), anche la successione $(\xi_{T_{x_k}})$ converge puntualmente verso 0. Dal teorema di Lebesgue sulla convergenza dominata segue dunque la relazione

$$\lim_k \int_{\{\xi_0=0, T_{x_k}<\infty\}} \phi(x_k + \xi_{T_{x_k}}) dP = \phi(z),$$

che, grazie a (7.3.9), si può anche scrivere nella forma

$$\lim_k H_D\phi(x_k) = \phi(z).$$

Il teorema è così dimostrato.

Il teorema seguente fornisce una condizione sufficiente ad assicurare che un assegnato punto di frontiera di D sia, per l'insieme D , un punto di uscita regolare.

(7.3.10) Teorema. *Affinché un punto x di ∂D appartenga a $\partial_r D$, è sufficiente che esistano un insieme boreliano K di \mathbb{R}^n , non trascurabile secondo Lebesgue, una successione (c_k) di numeri reali strettamente positivi, convergente verso 0, ed una successione (ρ_k) di rotazioni di \mathbb{R}^n intorno all'origine, tali che si abbia*

$$\bigcup_k (x + c_k \rho_k(K)) \subset D^c.$$

Dimostrazione. Senza ledere la generalità si può supporre che sia $x = 0$. Si ponga

$$t_k = c_k^2, \quad A_k = \{\xi_{t_k} \in c_k \rho_k(K)\} = \{\xi_{t_k} / \sqrt{t_k} \in \rho_k(K)\}.$$

Si ha allora

$$\limsup_k A_k \subset \limsup_k \{\xi_{t_k} \in D^c\} \subset \{T = 0\}.$$

Inoltre gli eventi A_n hanno, secondo P , una medesima probabilità non nulla. Ne segue $P\{T = 0\} > 0$, ossia, $P\{T = 0\} = 1$.

(7.3.11) Osservazione. Se un punto x della frontiera di D verifica la classica “condizione del cono” (cioè se esiste un cono non degenero, di vertice x , la cui intersezione con un opportuno intorno di x sia contenuta in D^c), allora x verifica banalmente la condizione del criterio precedente. Di conseguenza, affinché la funzione $H_D \phi$ sia soluzione del problema di Dirichlet (7.3.1), basta che D verifichi, in ogni suo punto della frontiera, la proprietà del cono. In particolare, quando D è un dominio convesso, esso verifica sempre la proprietà del cono (si può persino prendere come cono un semipiano intero) e quindi il problema di Dirichlet ha sempre soluzione per ogni funzione reale ϕ , continua su ∂D .

A

Richiami di Analisi e Probabilità

A.1 Spazi di Riesz e classi monotone

Supponiamo fissato un insieme non vuoto E . Si denota con \mathbb{R}^E lo spazio vettoriale ordinato costituito da tutte le funzioni da E in \mathbb{R} .

(A.1.1) Definizione. Una parte \mathcal{R} di \mathbb{R}^E sarà detta uno *spazio di Riesz* (o *spazio vettoriale reticolato*) se è un sottospazio vettoriale di \mathbb{R}^E stabile per le due “operazioni reticolari”:

$$(f, g) \mapsto f \wedge g, \quad (f, g) \mapsto f \vee g.$$

Ricordiamo che, data una funzione f definita su E , la *parte positiva* di f è la funzione $f^+ = f \vee 0$ mentre la *parte negativa* di f è la parte positiva di $-f$, ossia la funzione $f^- = (-f) \vee 0$.

Ricordiamo le seguenti formule (valide in ogni spazio di Riesz):

$$(A.1.2) \quad f = f^+ - f^-, \quad |f| = f^+ + f^-,$$

$$(A.1.3) \quad f \vee g = f + (g - f)^+ = \frac{1}{2}(f + g + |g - f|),$$

$$(A.1.4) \quad f \wedge g = -[(-f) \vee (-g)], \quad f^- = (-f)^+.$$

Utilizzando queste relazioni si riconosce facilmente che, affinché un sottospazio vettoriale di \mathbb{R}^E sia uno spazio di Riesz, è necessario e sufficiente ch’esso sia stabile per una qualsiasi delle tre operazioni seguenti:

$$f \mapsto f^+, \quad f \mapsto f^-, \quad f \mapsto |f|.$$

(A.1.5) Definizione. Quando uno spazio di Riesz \mathcal{R} , su un insieme E , è stabile per l’operazione $f \mapsto f \wedge 1$, si dice ch’esso possiede la *proprietà di Stone*.

Vale il risultato seguente, la cui dimostrazione si trova in G. Letta [31].

(A.1.6) Proposizione. Sia \mathcal{I} un insieme di parti di E , stabile per intersezione binaria, e si denoti con \mathcal{R} lo spazio vettoriale generato da $\{I_A : A \in \mathcal{I}\}$. Allora \mathcal{R} è uno spazio di Riesz verificante la proprietà di Stone.

(A.1.7) Definizione. Una parte \mathcal{M} di \mathbb{R}^E si dice *monotona* (o anche una *classe monotona*) se essa è stabile per la “convergenza dominata delle successioni monotone”, nel senso che essa possiede le proprietà seguenti:

(a) Per ogni successione crescente (f_n) di elementi di \mathcal{M} , maggiorati da un medesimo elemento di \mathcal{M} , si ha $\sup_n f_n \in \mathcal{M}$.

(b) Per ogni successione decrescente (f_n) di elementi di \mathcal{M} , minorati da un medesimo elemento di \mathcal{M} , si ha $\inf_n f_n \in \mathcal{M}$.

(A.1.8) Osservazione. Se \mathcal{M} è una classe monotona stabile per le operazioni reticolari, essa è *numerabilmente reticolata*, nel senso che contiene l’involuppo superiore (risp. inferiore) di una qualsiasi successione (f_n) di elementi di \mathcal{M} , tutti maggiorati (risp. minorati) da un medesimo elemento di \mathcal{M} . Ciò discende dalle ovvie relazioni:

$$\sup_n f_n = \sup_n (f_1 \vee \cdots \vee f_n), \quad \inf_n f_n = \inf_n (f_1 \wedge \cdots \wedge f_n).$$

(A.1.9) Definizione. Data una classe \mathcal{H} di funzioni reali definite su E , la minima tra tutte le parti monotone di \mathbb{R}^E contenenti \mathcal{H} si chiama *classe monotona generata* da \mathcal{H} .

(A.1.10) Osservazione. Si osservi che, se \mathcal{M} è una classe monotona generata da \mathcal{H} , ogni elemento di \mathcal{M} è compreso tra due elementi di \mathcal{H} . In effetti, la classe delle funzioni che possiedono questa proprietà è monotona e contiene \mathcal{H} , dunque contiene \mathcal{M} .

Utilizzando la definizione di classe monotona generata, è facile verificare la proposizione seguente:

(A.1.11) Proposizione. Se \mathcal{H} è un sottospazio vettoriale (risp. spazio di Riesz) di \mathbb{R}^E , tale è la classe monotona \mathcal{M} generata da \mathcal{H} .

Enunciamo di seguito i quattro *teoremi delle classi monotone*, per una dimostrazione si veda G. Letta [31].

(A.1.12) Teorema (elementare delle classi monotone). Sia \mathcal{M} uno spazio di Riesz monotono, contenente le costanti e costituito da funzioni limitate, definite sull’insieme E . Si denoti con \mathcal{E} la classe costituita da tutte le parti di E la cui funzione indicatrice appartenga a \mathcal{M} . Allora:

(a) \mathcal{E} è una tribù su E .

(b) \mathcal{M} coincide con la classe di tutte le funzioni limitate e misurabili su (E, \mathcal{E}) .

(c) \mathcal{E} coincide con la tribù $\mathcal{T}(\mathcal{M})$ generata da \mathcal{M} (ossia la minima tribù che rende misurabili tutte le funzioni di \mathcal{M}).

(A.1.13) Teorema (delle classi monotone). Siano dati uno spazio misurabile (E, \mathcal{E}) e una classe \mathcal{H} di funzioni limitate su E , dotate delle proprietà seguenti:

(a) La tribù $\mathcal{T}(\mathcal{H})$ generata da \mathcal{H} coincide con l'intera tribù \mathcal{E} .

(b) \mathcal{H} contiene la costante 1 ed è stabile per la moltiplicazione, ossia per l'operazione $(f, g) \mapsto fg$.

Allora ogni spazio vettoriale monotono \mathcal{L} (di funzioni reali su E), il quale contenga \mathcal{H} , contiene la classe di tutte le funzioni limitate e misurabili su (E, \mathcal{E}) .

(A.1.14) Teorema (secondo, delle classi monotone). Nello spazio misurabile (E, \mathcal{E}) , sia assegnato un sistema di generatori \mathcal{I} per la tribù \mathcal{E} , stabile per l'operazione di intersezione di due insiemi. Sia poi \mathcal{L} un sottospazio vettoriale e monotono di \mathbb{R}^E , contenente le indicatori degli insiemi di \mathcal{I} . Allora \mathcal{L} contiene ogni funzione della forma

$$fI_{A_1 \cup \dots \cup A_n},$$

con f funzione misurabile e limitata su (E, \mathcal{E}) e (A_i) n -upla di elementi di \mathcal{I} . (In altri termini: \mathcal{L} contiene ogni funzione misurabile e limitata su (E, \mathcal{E}) che sia nulla fuori di una riunione numerabile di elementi di \mathcal{I}).

(A.1.15) Teorema (terzo, delle classi monotone). Siano assegnati uno spazio misurabile (E, \mathcal{E}) e uno spazio di Riesz \mathcal{R} (di funzioni reali definite su E) il quale possieda la proprietà di Stone e generi la tribù \mathcal{E} . Sia poi \mathcal{L} una parte monotona di \mathbb{R}^E contenente \mathcal{R} . Allora \mathcal{L} contiene ogni funzione misurabile su (E, \mathcal{E}) che sia compresa tra due funzioni di \mathcal{R} .

Come corollario del teorema delle classi monotone (A.1.13), si prova facilmente il risultato seguente:

(A.1.16) Lemma di misurabilità di Doob. Dati due spazi misurabili (E, \mathcal{E}) e (F, \mathcal{F}) , sia f un'applicazione da E a F , tale che la tribù \mathcal{E} sia la minima che rende misurabile la funzione f . Allora, ogni funzione numerica g , misurabile su (E, \mathcal{E}) , è della forma $h \circ f$, con h opportuna funzione numerica e misurabile su (F, \mathcal{F}) .

Similmente, tramite il teorema delle classi monotone (A.1.14), deduce il criterio seguente (per la cui dimostrazione si veda G. Letta [31]).

(A.1.17) Criterio per la coincidenza di due misure. Dato uno spazio misurabile (E, \mathcal{E}) , sia \mathcal{I} un insieme di parti di E che generi la tribù \mathcal{E} , stabile per l'operazione di intersezione binaria, e tale che l'intero spazio sia unione numerabile di elementi di \mathcal{I} . Siano μ, ν due misure su (E, \mathcal{E}) tali che si abbia

$$\mu(A) = \nu(A) < \infty \quad \text{per ogni } A \in \mathcal{I}.$$

Allora risulta $\mu = \nu$.

A.2 La nozione di speranza condizionale

Supponiamo fissato uno spazio probabilizzato (Ω, \mathcal{A}, P) , e supponiamo fissata una sottotribù \mathcal{F} di \mathcal{A} . Si denoti con Q la restrizione di P alla sottotribù \mathcal{F} .

(A.2.1) Definizione. Data una variabile aleatoria reale integrabile X sullo spazio probabilizzato (Ω, \mathcal{A}, P) , si dice che una variabile aleatoria reale V , misurabile secondo \mathcal{F} , e integrabile secondo Q , è una *versione della speranza condizionale* di X rispetto a \mathcal{F} , se risulta

$$\int_A X \, dP = \int_A V \, dP \quad \text{per ogni } A \in \mathcal{F}.$$

(A.2.2) Teorema. Sia X una variabile aleatoria reale su (Ω, \mathcal{A}, P) , integrabile secondo P . Allora esiste sempre un'unica (a meno di equivalenza modulo Q) versione della speranza condizionale di X rispetto ad una sottotribù \mathcal{F} di \mathcal{A} . In particolare si può scegliere una versione positiva se tale è la variabile aleatoria X .

Dimostrazione. Basta dimostrare il teorema nel caso in cui X è positiva. In generale, infatti, basterà ragionare separatamente su X^+ e X^- . Osserviamo che la misura μ , definita da

$$(A.2.3) \quad \mu(A) = \int_A X \, dP \quad \text{per ogni } A \in \mathcal{A},$$

è una misura assolutamente continua rispetto a P perché ha densità X secondo P . Si denoti con ν la restrizione di μ alla sottotribù \mathcal{F} . Si riconosce facilmente che la misura ν è assolutamente continua rispetto a Q , e dunque, per il teorema di Radon-Nikodým, esiste una densità V di ν rispetto a Q . Ne segue l'uguaglianza

$$\int_A X \, dP = \int_A V \, dQ \quad \text{per ogni } A \in \mathcal{F}.$$

Per concludere basta osservare che, se U è una variabile aleatoria misurabile rispetto alla sottotribù \mathcal{F} , allora essa è integrabile secondo Q se, e soltanto se, lo è secondo P , e quando queste condizioni equivalenti sono verificate, si ha

$$\int U \, dQ = \int U \, dP.$$

Per quanto riguarda l'unicità, osserviamo che, se V_1, V_2 sono due versioni della speranza condizionale di X rispetto a \mathcal{F} , allora le misure

$$\nu_1(A) = \int_A V_1 \, dP, \quad \nu_2(A) = \int_A V_2 \, dP$$

coincidono sulla tribù \mathcal{F} essendo eguali alla misura ν : le due densità devono dunque essere necessariamente equivalenti modulo Q .

(A.2.4) Definizione. Nelle stesse ipotesi di (A.2.1), si chiama *speranza condizionale* di X rispetto a \mathcal{F} , e si denota con $P[X|\mathcal{F}]$, la classe d'equivalenza (in $L^1(Q)$) costituita da tutte le versioni della speranza condizionale di X rispetto a \mathcal{F} .

Riassumiamo, nella proposizione seguente, le proprietà elementari della speranza condizionale. Per una dimostrazione si rimanda all'appendice di B. Øksendal [37] oppure al libro di J. Jacod e P. Protter [27].

(A.2.5) Proposizione. Siano X, Y una coppia di variabili aleatorie reali e integrabili sullo spazio probabilizzato (Ω, \mathcal{A}, P) , e siano a, b una coppia di numeri reali. Allora:

- (a) si ha $P[aX + bY|\mathcal{F}] = aP[X|\mathcal{F}] + bP[Y|\mathcal{F}]$.
- (b) Se si denota con V una versione di $P[X|\mathcal{F}]$, allora $P[V] = P[X]$.
- (c) Se X è misurabile rispetto alla tribù \mathcal{F} , allora X è una versione di $P[X|\mathcal{F}]$.
- (d) Se X è indipendente dalla tribù \mathcal{F} , allora la costante $P[X]$ è una versione di $P[X|\mathcal{F}]$.
- (e) Se Y è misurabile rispetto alla tribù \mathcal{F} , allora $P[XY|\mathcal{F}] = YP[X|\mathcal{F}]$.

È facile provare, inoltre, le seguenti versioni condizionali delle proprietà dell'integrazione.

(A.2.6) Proposizione. Valgono i fatti seguenti:

(a) (*Teorema di Beppo Levi condizionale*) Sia X un elemento di $\mathcal{L}^1(P)$, e sia (X_n) una successione crescente di elementi di $\mathcal{L}^1(P)$ ammettenti X come inviluppo superiore. Sussiste allora la relazione

$$P[X|\mathcal{F}] = \sup_n P[X_n|\mathcal{F}].$$

(b) (*Lemma di Fatou condizionale*) Sia X un elemento di $\mathcal{L}^1(P)$, e sia (X_n) una successione di elementi di $\mathcal{L}^1(P)$ tutti minorati da un medesimo elemento di $\mathcal{L}^1(P)$, e ammettenti X come limite inferiore. Sussiste allora la relazione

$$P[X|\mathcal{F}] \leq \liminf_n P[X_n|\mathcal{F}].$$

(c) (*Disuguaglianza di Jensen condizionale*) Sia X una variabile aleatoria integrabile, e sia φ una funzione reale convessa su \mathbb{R} . Si denoti con V una versione di $P[X|\mathcal{F}]$, e con W una versione di $P[\varphi(X)|\mathcal{F}]$. L'evento $\{\varphi(V) \leq W\}$ è allora quasi certo.

(d) (*Proprietà della tribù intermedia*) Sia \mathcal{G} una sottotribù di \mathcal{A} . Supponiamo che si abbia $\mathcal{F} \subset \mathcal{G}$. Sia inoltre X una variabile aleatoria reale integrabile, e si denoti con Y una versione di $P[X|\mathcal{G}]$. Si ha allora

$$P[X|\mathcal{F}] = P[Y|\mathcal{F}].$$

La proposizione seguente fornisce un utile strumento per il calcolo della speranza condizionale.

(A.2.7) Proposizione. Siano X, Y una coppia di variabili aleatorie a valori negli spazi misurabili $(E, \mathcal{E}), (F, \mathcal{F})$. Sia g una funzione reale, limitata e misurabile sullo spazio $(E \times F, \mathcal{E} \otimes \mathcal{F})$. Si supponga che X sia misurabile rispetto alla tribù \mathcal{F} e che Y sia indipendente da \mathcal{F} . Si ponga

$$G(x) = P[g(x, Y)].$$

La variabile aleatoria $G(X)$ è allora una versione di $P[g(X, Y) | \mathcal{F}]$.

Dimostrazione. Osserviamo innanzitutto che, se g è della forma $g_1 \otimes g_2$, con g_1, g_2 funzioni reali, misurabili e limitate su (E, \mathcal{E}) e (F, \mathcal{F}) rispettivamente, allora la tesi è ovvia perché si ha

$$P[g_1(X)g_2(Y) | \mathcal{F}] = g_1(X)P[g_2(Y) | \mathcal{F}] = g_1(X)P[g_2(Y)].$$

Ora, denotato con \mathcal{L} l'insieme costituito da tutte le funzioni reali e limitate g , misurabili su $(E \times F, \mathcal{E} \otimes \mathcal{F})$ e verificanti la proprietà desiderata, osserviamo che si tratta di uno spazio vettoriale monotono che, per giunta, contiene la classe \mathcal{H} costituita da tutte le funzioni g della forma $g_1 \otimes g_2$, con g_1, g_2 opportune funzioni reali e limitate su (E, \mathcal{E}) e (F, \mathcal{F}) rispettivamente. La tesi segue dunque per il teorema delle classi monotone (A.1.13).

A.3 La nozione di nucleo

Siano dati due spazi misurabili $(E, \mathcal{E}), (F, \mathcal{F})$ e una famiglia

$$(A.3.1) \quad N = (N(x, \cdot))_{x \in E}$$

di misure su \mathcal{F} , avente E come insieme degli indici. In queste condizioni, se g è una funzione misurabile positiva su (F, \mathcal{F}) , denoteremo con Ng la funzione così definita su E :

$$(A.3.2) \quad Ng(x) = \langle N(x, \cdot), g \rangle = \int N(x, dy) g(y).$$

Quando g sia l'indicatrice di un elemento B di \mathcal{F} , avremo dunque, in particolare,

$$(A.3.3) \quad NI_B(x) = N(x, B).$$

Osserviamo che, per ogni successione (g_n) di funzioni positive e misurabili su (F, \mathcal{F}) , valgono le due proprietà seguenti:

$$(A.3.4) \quad g_n \uparrow g \Rightarrow Ng_n \uparrow Ng, \quad g = \sum_n g_n \Rightarrow Ng = \sum_n Ng_n.$$

(A.3.5) Definizione. Nelle ipotesi sopra precisate, si dice che la famiglia di misure (A.3.1) è un *nucleo*, relativo alla coppia di spazi misurabili $(E, \mathcal{E}), (F, \mathcal{F})$ (o anche un nucleo *da* (E, \mathcal{E}) *a* (F, \mathcal{F})) se, per ogni funzione g misurabile e positiva su (F, \mathcal{F}) , la funzione positiva Ng definita su E da (A.3.2) è misurabile su (E, \mathcal{E}) . Si dice, poi, che un nucleo N è *markoviano* se, per ogni elemento x in E , la misura $N(x, \cdot)$ è di probabilità.

Osserviamo che se N è un nucleo markoviano, tutte le proprietà dei nuclei che coinvolgono funzioni misurabili e positive si possono estendere a funzioni misurabili e limitate ponendo $Nf = Nf^+ - Nf^-$.

(A.3.6) Osservazione. Nelle ipotesi della definizione precedente, è immediato riconoscere che, affinché la famiglia di misure (A.3.1) sia un nucleo è (necessario e) sufficiente che, per ogni elemento B di \mathcal{F} , la funzione NI_B definita da (A.3.3) sia misurabile su (E, \mathcal{E}) . Infatti, supposto che questa condizione sia soddisfatta, si riconosce innanzitutto facilmente che, per ogni funzione g positiva e semplice su (F, \mathcal{F}) , la funzione Ng è misurabile su (E, \mathcal{E}) . In seguito si estende questo risultato al caso di un'arbitraria funzione misurabile positiva g , esprimendo g come inviluppo superiore di una successione crescente (g_n) di funzioni semplici positive, e riconducendosi al caso precedente con l'ausilio della proprietà (A.3.4).

(A.3.7) Proposizione. Sia N un nucleo da (E, \mathcal{E}) a (F, \mathcal{F}) e sia μ una misura su (E, \mathcal{E}) . Allora:

(a) La funzione μN che ad ogni elemento B di \mathcal{F} associa il numero $\mu N(B)$ definito da

$$(A.3.8) \quad \mu N(B) = \langle \mu, NI_B \rangle = \int \mu(dx) N(x, B)$$

è una misura su \mathcal{F} .

(b) Per ogni funzione g misurabile e positiva su (F, \mathcal{F}) , sussiste la seguente “relazione di dualità”:

$$(A.3.9) \quad \langle \mu N, g \rangle = \langle \mu, Ng \rangle.$$

Dimostrazione. Per ogni elemento B di \mathcal{F} e ogni successione (B_n) di elementi di \mathcal{F} , a due a due disgiunti, la cui unione sia eguale a B , l'ovvia relazione $I_B = \sum_n I_{B_n}$ implica

$$NI_B = \sum_n NI_{B_n}$$

e quindi $\langle \mu N, I_B \rangle = \sum_n \langle \mu, NI_{B_n} \rangle$, ossia $\mu N(B) = \sum_n \mu N(B_n)$. Ciò prova che μN è una misura su \mathcal{F} .

Proviamo ora l'affermazione (b). Essa è immediata nel caso particolare in cui g sia la funzione indicatrice di un elemento B di \mathcal{F} . In questo caso, essa discende infatti direttamente dalla relazione (A.3.8) che definisce la misura μN . Immediata è poi l'estensione al caso in cui g sia semplice. Infine, nel caso generale,

dopo aver espresso g come inviluppo superiore di una successione crescente (g_n) di funzioni semplici positive, basta ricondursi al caso precedente servendosi della proprietà (A.3.4) e del teorema di Beppo Levi.

Come risulta dalla proposizione precedente, un nucleo N da (E, \mathcal{E}) a (F, \mathcal{F}) induce due distinte operazioni:

- un'operazione $\mu \mapsto \mu N$ che trasforma ogni misura μ su \mathcal{E} in una misura μN su \mathcal{F} ;
- un'operazione $g \mapsto Ng$ che trasforma ogni funzione g positiva e misurabile su (F, \mathcal{F}) in una funzione Ng positiva e misurabile su (E, \mathcal{E}) .

Queste due operazioni sono tra loro legate dalla fondamentale relazione di dualità (A.3.9). Si osservi che la seconda di esse individua completamente il nucleo, grazie alla relazione $N(x, B) = NI_B(x)$. Anche la prima operazione individua il nucleo: si ha infatti $N(x, \cdot) = \varepsilon_x N$, dove ε_x designa (per ogni x in E) la misura così definita su \mathcal{E} :

$$\varepsilon_x(A) = I_A(x) \quad \text{per ogni } A \text{ di } \mathcal{E}$$

(misura di Dirac definita da una massa unitaria concentrata nel punto x).

(A.3.10) Osservazione. Siano ora assegnati un nucleo N da (E, \mathcal{E}) a (F, \mathcal{F}) e un nucleo Q da (F, \mathcal{F}) a un ulteriore spazio misurabile (G, \mathcal{G}) . È facile allora riconoscere che l'applicazione $g \mapsto NQg$ (ottenuta componendo la trasformazione indotta da Q con quella indotta da N) possiede ancora le proprietà (a), (b), sicché è a sua volta indotta da un nucleo (relativo alla coppia di spazi $(E, \mathcal{E}), (G, \mathcal{G})$). Il nucleo così caratterizzato si denota con NQ e si chiama *nucleo composto* ottenuto componendo i due nuclei N, Q . Si riconosce immediatamente che, per ogni misura μ su (E, \mathcal{E}) , la trasformata di μ mediante il nucleo NQ è data da

$$\mu(NQ) = (\mu N)Q,$$

ossia coincide con la trasformata, mediante Q , della trasformata di μ mediante N .

La proposizione che segue fornisce un criterio (molto utile) per decidere se un'assegnata famiglia di misure sia un nucleo.

(A.3.11) Proposizione. Siano dati due spazi misurabili $(E, \mathcal{E}), (F, \mathcal{F})$ e una famiglia $N = (N(x, \cdot))_{x \in E}$ di misure su \mathcal{F} , avente E come insieme degli indici. Sia \mathcal{J} una base per la tribù \mathcal{F} . Si supponga che, per ogni elemento B di \mathcal{J} , la funzione NI_B sia finita e misurabile su (E, \mathcal{E}) . Allora N è un nucleo, relativo alla coppia $(E, \mathcal{E}), (F, \mathcal{F})$.

Dimostrazione. Denotiamo con \mathcal{L} la classe costituita da tutte le funzioni g , appartenenti a $\bigcap_{x \in E} \mathcal{L}^1(N(x, \cdot))$, tali che la funzione Ng sia misurabile su (E, \mathcal{E}) . Allora \mathcal{L} è uno spazio vettoriale monotono e contiene le indicatrici degli insiemi della classe \mathcal{J} . Ne segue, grazie a (A.1.14), che \mathcal{L} contiene l'indicatrice di ogni insieme del tipo $B \cap (B_1 \cup \dots \cup B_n)$, con B elemento di \mathcal{F} e $n \geq 1$. In altri termini: per ogni

elemento B di \mathcal{F} , ciascuna delle funzioni $NI_{B \cap (B_1 \cup \dots \cup B_n)}$ è misurabile su (E, \mathcal{E}) . Tale è dunque anche la funzione

$$NI_B = \sup_n NI_{B \cap (B_1 \cup \dots \cup B_n)}.$$

Ciò prova che N è un nucleo.

A.4 Il teorema delle contrazioni

Supponiamo fissato uno spazio metrico e completo E . Si denoti con d una sua distanza.

(A.4.1) Definizione. Una funzione f da E in sé si dice una *contrazione* su E se esiste una costante reale strettamente positiva α inferiore a 1, tale che si abbia

$$d(f(x), f(y)) \leq \alpha d(x, y) \quad \text{per ogni } x, y \in E.$$

Il seguente è un classico risultato di Analisi reale. Una dimostrazione (ed una esposizione approfondita di questo argomento) si può trovare nel testo di M. Abate [1].

(A.4.2) Teorema (delle contrazioni). Sia f una contrazione su E . Allora esiste un unico punto x in E , tale che si abbia $f(x) = x$.

Più in generale, quando la contrazione f abbia un parametro t , vale la seguente estensione del teorema precedente (di semplice dimostrazione).

(A.4.3) Teorema (delle contrazioni con parametro). Sia F uno spazio metrico. Si denoti con d' una distanza su F . Sia f una funzione da $E \times F$ in E che verifichi le condizioni seguenti:

- (a) Per ogni elemento t in F , la funzione $x \mapsto f(x, t)$ è una contrazione su E .
- (b) Per ogni elemento x di E , la funzione $t \mapsto f(x, t)$ è lipschitziana di costante L , vale cioè la relazione

$$d(f(x, t), f(x, s)) \leq L d'(s, t) \quad \text{per ogni } s, t \in F.$$

Per ogni elemento t in F esiste allora un unico elemento x_t in E , tale che si abbia $f(x_t, t) = x_t$. Inoltre la funzione $t \mapsto x_t$, di F in E è lipschitziana di costante $L/(1 - \alpha)$.

Infine ricordiamo che, data una contrazione f su E , è possibile scrivere esplicitamente una successione di elementi di E la quale converga verso il punto fisso. Più precisamente.

(A.4.4) Corollario. *Nelle stesse ipotesi del teorema delle contrazioni, la successione (x_n) definita per ricorrenza da*

$$x_0 \in E, \quad x_{n+1} = f(x_n),$$

converge (in E) verso il punto fisso x di f .

A.5 Funzioni armoniche

Sia D un aperto in \mathbb{R}^n . Si dice che una funzione u di D in \mathbb{R} è *armonica* (su D) se è di classe \mathcal{C}^2 e se verifica $\Delta u = 0$, dove Δ denota l'operatore di Laplace definito da

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}.$$

Richiamiamo le seguenti proprietà delle funzioni armoniche. Per ulteriori approfondimenti su questo argomento, e per le dimostrazioni di questi teoremi, si rimanda al classico testo di D. Gilbarg e N.S. Trudinger [20] oppure al *Quaderno* di E. Giusti [21].

(A.5.1) Teorema. *Sia u una funzione continua su \overline{D} . Le condizioni che seguono sono allora equivalenti:*

- (a) u è armonica.
- (b) u verifica la cosiddetta *proprietà della media sulle sfere*, nel senso che, per ogni palla B , di centro x , che sia contenuta in D , risulta

$$u(x) = \int_{\partial B} u(y) \lambda(dy),$$

dove λ denota la ripartizione uniforme su ∂B (ossia l'unica misura di probabilità sulla tribù boreliana di ∂B che sia invariante per rotazioni attorno al centro).

(A.5.2) Teorema (principio del massimo). *Sia u una funzione reale, definita e continua su \overline{D} e di classe \mathcal{C}^2 su D . Supponiamo che u sia armonica in D , e che essa assuma valore massimo in D . Allora u è costante.*

Bibliografia

- [1] M. ABATE. *An introduction to hyperbolic dynamical systems*. Istituti poligrafici internazionali, Pisa-Roma (2001)
- [2] P. ACQUISTAPACE. *Appunti di Analisi convessa*. Pisa (2004), attualmente disponibili all'indirizzo <http://www.dm.unipi.it/~acquistp>
- [3] P. ACQUISTAPACE. *Appunti di Analisi funzionale*. Pisa (2004), attualmente disponibili all'indirizzo <http://www.dm.unipi.it/~acquistp>
- [4] P. BALDI. *Equazioni differenziali stocastiche e applicazioni*. Quaderni U.M.I., Pitagora, Bologna (2000)
- [5] P. BILLINGSLEY. *Convergence of probability measures*. Wiley, New York (1999)
- [6] P. BILLINGSLEY. *Probability and measure*. Wiley, New York (1995)
- [7] N. BOURBAKI. *Éléments de mathématique, topologie générale. Chap. I à IV*. Hermann, Paris (1971)
- [8] N. BOURBAKI. *Éléments de mathématique, topologie générale. Chap. V à X*. Hermann, Paris (1974)
- [9] S. CERRAI. *Second order PDEs in finite and infinite dimension. A probabilistic approach*. Springer LN 1762, Berlin (2000)
- [10] K.L. CHUNG. *A course in probability theory*. Academic Press, New York (1968)
- [11] K.L. CHUNG. *Green, Brown and Probability & Brownian Motion on the line*. World Scientific, Singapore (2002)
- [12] G. DA PRATO. *An introduction to Infinite Dimensional Analysis*. Appunti SNS, Pisa (2001)
- [13] G. DA PRATO. *Introduction to differential stochastic equations*. Appunti SNS, Pisa (1998)
- [14] C. DELLACHERIE, P.-A. MEYER. *Probabilités et potentiel, Chap. I à IV*. Hermann, Paris (1975)

- [15] C. DELLACHERIE, P.-A. MEYER. *Probabilités et potentiel, Chap. V à VIII*. Hermann, Paris (1980)
- [16] C. DELLACHERIE, P.-A. MEYER. *Probabilités et potentiel, Chap XII à XVI*. Hermann, Paris (1987)
- [17] J.L. DOOB. *Stochastic processes*. Wiley, New York (1953)
- [18] D. FOATA, A. FUCHS. *Calcul des probabilités*. Dunod, Paris (1998)
- [19] D. FOATA, A. FUCHS. *Processus stochastiques*. Dunod, Paris (2002)
- [20] D. GILBARG, N.S. TRUDINGER. *Elliptic partial differential equations of second order*. Springer, Berlin (1977)
- [21] E. GIUSTI. *Equazioni ellittiche del secondo ordine*. Quaderni U.M.I., Pitagora, Bologna (1978)
- [22] P.R. HALMOS. *Lectures on Boolean Algebras*. Springer, Berlin (1963)
- [23] K. ITÔ. *On a formula concerning stochastic differentials*. Nagoya Math. J., **3** (1951), pp. 55-65
- [24] K. ITÔ. *On a stochastic integral equation*. Proc. Japan. Acad. Ser. A Math. Sci., **22** (1946), pp. 32-35
- [25] K. ITÔ. *On stochastic differential equations*. Mem. Amer. Math. Soc., **4** (1951)
- [26] K. ITÔ. *Stochastic integral*. Proc. Imp. Acad. Tokyo, **20** (1944), pp. 519-524
- [27] J. JACOD, P. PROTTER. *Probability essentials*. Universitext, Springer, Berlin (2004)
- [28] I. KARATZAS, S.E. SHREVE. *Brownian motion and stochastic calculus*. GTM, Springer, Berlin (1988)
- [29] D. LAMBERTON, B. LAPEYRE. *Introduction au calcul stochastique appliqué à la finance*. Ellipses, Paris (1997)
- [30] G. LETTA. *Integration stochastique*. Bollettino U.M.I., (6), **2-A** (1983), pp. 1-20
- [31] G. LETTA. *Martingales et integration stochastique*. Quaderni SNS, Pisa (1984)
- [32] G. LETTA. *Probabilità elementare*. Zanichelli (1993)
- [33] G. LETTA. *Teoria elementare dell'integrazione*. Boringhieri, Torino (1976)
- [34] M.E. MANCINO. *Dal moto browniano all'equazione di Schrödinger*. Tesi di laurea, Pisa (1990)
- [35] P.-A. MEYER. *Caractérisation des semimartingales d'après Dellacherie*. Sémin. Prob. XIII, Springer LN 721, Berlin (1979), pp. 620-623
- [36] P.-A. MEYER. *Processus de Markov*. Springer LN 26, Berlin (1976)

- [37] B. ØKSENDAL. *Stochastic differential equations*. Universitext, Springer, Berlin (2003)
- [38] N. PINTACUDA. *Probabilità*. Decibel, Padova (1994)
- [39] P. PROTTER. *Stochastic integration and differential equations*. Springer, Berlin (2003)
- [40] D. REVUZ, M. YOR. *Continuous martingales and brownian motion*. Springer, Berlin (1991)
- [41] J. ZABCZYK. *Topics in Stochastic Processes*. Quaderni SNS, Pisa (2004)